PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:
Donald W. Kufe

Serial No.: 10/733,212

Filed: December 11, 2003

For: REGULATION OF CELL GROWTH BY MUC1

Confirmation No. 7998

Group Art Unit: 1633

Examiner: Hill, Kevin Kai

Atty. Dkt. No.: GENU:009USD1

CERTIFICATE OF ELECTRONIC TRANSMISSION
37 C.F.R. § 1.8

I hereby certify that this correspondence is being electronically filed with the United States Patent and Trademark Office via EFS-Web on the date below:

August 25, 2009
Date

Monica A. De La Paz

APPEAL BRIEF

TABLE OF CONTENTS

Page

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:
Donald W. Kufe

Serial No.: 10/733,212

Filed: December 11, 2003

For: REGULATION OF CELL GROWTH BY MUC1

Confirmation No. 7998

Group Art Unit: 1633

Examiner: Hill, Kevin Kai

Atty. Dkt. No.: GENU:009USD1

APPEAL BRIEF

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-01450

Dear Sir:

This Appeal Brief is filed in response to the final Office Action mailed on March 2, 2009, and the Advisory Action mailed on May 28, 2009. A Notice of Appeal was filed on July 1, 2009, making this brief due on September 1, 2009. Also included herewith is the fee for the brief. No other fees are believed due in connection with this filing; however, should appellants' payment be missing or deficient, or should any fees be due, the Commissioner is authorized to debit Fulbright & Jaworski L.L.P. Deposit Acct. No. 50-1212/GENU:009USD1/SLH.

I. Real Party in Interest

The real parties in interest are the assignee, the Dana Farber Cancer Institute, Boston, MA, and the licensee, Genus Oncology, LLC.

II. Related Appeals and Interferences

There are no related appeals or interferences.

III. Status of the Claims

Claims 1-56 were filed with the original application. The claims were subjected to a restriction requirement, and as a result, claims 5-10, 16-18, 20-36, 39, 41, 43, 44, 48-51 and 53-56 stand withdrawn. Thus, claims 1-4, 11-15, 19 37, 38, 40, 42, 45-47 and 52 were examined. Claims 57 was subsequently added, and claims 2-11, 16-18, 20-36, 39, 41, 43, 44, 48-51 and 43-57 were canceled. Thus, claims 1, 12-15, 189, 37, 38, 40, 42, 45-47 and 52 are pending, under examination, stand rejected and are appealed. The pending claims are attached in Appendix A.

IV. Status of the Amendments

The amendments offered following mailing the final Office Action were entered pursuant to the Advisory Action of May 28, 2009.

V. Summary of the Claimed Subject Matter

Independent claim 1 is supported in the specification, for example, at page 1, line 24 to page 2, line 1, and page 17, lines 28-30.

VI. Grounds of Rejection to be Reviewed on Appeal

Are claims 1, 5, 7-9, 13, 15-17 and 22-26 obvious over Li *et al.* (1998; Exhibit 1) in view of Yamamoto *et al.* (1997; Exhibit 2) and Barker *et al.* (U.S. Patent 5, 851,775; Exhibit 3) as evidenced by Zrihan-Licht *et al.* (1994; Exhibit 4) under 35 U.S.C. §103?

VII. Argument

A. Standard of Review

Findings of fact and conclusions of law by the U.S. Patent and Trademark Office must be made in accordance with the Administrative Procedure Act, 5 U.S.C. §706(A), (E), 1994. *Dickinson v. Zurko*, 527 U.S. 150, 158 (1999). Moreover, the Federal Circuit has held that findings of fact by the Board of Patent Appeals and Interferences must be supported by "substantial evidence" within the record. *In re Gartside*, 203 F.3d 1305, 1315 (Fed. Cir. 2000). In *In re Gartside*, the Federal Circuit stated that "the 'substantial evidence' standard asks whether a reasonable fact finder could have arrived at the agency's decision." *Id.* at 1312. Accordingly, it necessarily follows that an examiner's position on appeal must be supported by "substantial evidence" within the record in order to be upheld by the Board of Patent Appeals and Interferences.

B. Rejection Under 35 U.S.C. §103

Claims 1, 5, 7-9, 13, 15-17 and 22-26 are rejected as obvious over Li *et al.* (1998; Exhibit 1) in view of Yamamoto *et al.* (1997; Exhibit 2) and Barker *et al.* (U.S. Patent 5, 851,775; Exhibit 3) as evidenced by Zrihan-Licht *et al.* (1994; Exhibit 4). The examiner cites Li and Yamamoto as providing methods of identifying a compound that inhibits binding of the β-catenin tumor progressor to a MUC1 test site. Barker is said to provide motivation for the use of a peptide fragment of β-catenin, and Zrihan-Licht is said to teach that the MUC1 test agent will necessarily be phosphorylated at the YEKV site. Appellant traverses.

i. The Examiner's Burden

In rejecting claims under 35 U.S.C. §103, the examiner bears the initial burden of presenting a *prima facie* case of obviousness. See *In re Rijckaert*, 9 F.3d 1531, 1532, 28 USPQ2d 1955, 1956 (Fed. Cir. 1993). A finding of obviousness requires that "the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains." 35 U.S.C. §103(a). In setting forth a *prima facie* case of obviousness, it is necessary to show "some articulated reasoning with some rational underpinning to support the legal conclusion of obviousness." *KSR International Co. v. Teleflex Inc.*, 127 S.Ct. 1727, 82 U.S.P.Q.2d 1385 (2007) (quoting *In re Kahn*, 441 F.3d 977, 988 (Fed. Cir. 2006)).

ii. Appellants' Position

In the present case, there is *no prima* facie case of obviousness for the following reasons. Li teaches that glycogen synthase kinase 3β binds to an STDRSPYE site in MUC1 and phosphorylates the serine that is adjacent to the proline. This phosphorylation decreases the binding of MUC1 to β-catenin. Li does not teach or suggest that phosphorylation of a YEKV site increases binding of MUC1 to β-catenin. The examiner has cited to FIG. 5 of Li as teaching GSK3β as the test agent. However, there is no information in Li to teach or suggest that the test agent in Li was phosphorylated at a YEKV site. Nor would this be inherent, as it is possible for a YEKV site to not be phosphorylated, and Li teaches that it is phosphorylation of a serine residue that affects interaction of MUC1 with β-catenin, not YEKV.

Yamamoto does not provide any teaching or suggestion concerning a MUC1 test agent phosphorylated at a YEKV site. Rather, it concerns certain studies demonstrating that DF3

(MUC1) binds directly to β-catenin and that the SXXXXXSSL motif in DF3 is responsible for this interaction.

Further, as admitted by the examiner, neither Li nor Yamamoto teach that the β-catenin test agent is a peptide fragment. Barker is cited as teaching that certain assays may be conducted utilizing a β-catenin fragment that is shorter than the full-length tumor progressor. It is not cited as providing any teaching or suggestion concerning assays concerning any MUC1 test agent, much less one that is phosphorylated at a YEKV site. The examiner admits that neither Li, Yamamoto, nor Barker teach that the MUC1 test agent includes a phosphorylated YEKV site. *See* Final Office Action, page 9.

While Zrihan-Licht discloses that MUC1 proteins are "extensively phosphorylated" and that phosphorylation occurs "primarily on tyrosine residues" (Abstract), it does not specifically teach phosphorylation of the YEKV site of MUC1. Indeed, the MUC1 protein includes 13 tyrosine residues, of which 7 are in the cytoplasmic domain, and there is no information in this reference or in any of the other references to suggest that the YEKV tyrosine residue, out of all of the amino acids of MUC1, is critical for binding to β-catenin. Further, Zrihan-Licht teaches that other residues may undergo phosphorylation, including serine residues. *See* p. 131, right col., third para. Still further, Zrihan-Licht teaches that the sequence YEEV is important for interaction with SH2 domain-containing tyrosine kinases, thus teaching away from the importance of a YEKV site. In addition, one of ordinary skill in the art would further be led away from the importance of phosphorylation of a YEKV site because, as discussed above, Li teaches that it is a serine residue that affects interaction of MUC1 with β-catenin and Yamamoto teaches that the SXXXXXSSL motif in DF3 is responsible for this interaction.

Thus, it is again submitted that there is no *prima facie* case of obviousness based on the combination of references cited by the examiner. There is no rationale that would have led one of ordinary skill in the art, at the time of the invention, to believe that the YEKV site of MUC1 is critical for binding to β-catenin, and thus a critical target for screening.

iii. The Examiner's Rebuttal Fails

In the Advisory Action mailed on May 28, 2009, the examiner found the preceding line of argument unpersuasive, and offered the following points in rebuttal.

First, it was argued that appellants were improperly addressing the references individually, and not as a whole. This is incorrect. Appellants were pointing out specific defects in the references, and the incorrect nature of the examiner's assumptions therefrom. When viewed in light of these critical deficiencies, the references cannot, even when taken as a whole, suggest the present invention. This is because they neither individually ***nor collectively*** provide any evidence that the YEKV motif is integral to β-catenin's interaction with MUC1.

Second, turning to Li, the examiner argues that the reference teaches that tyrosine residues flank the identified β-catenin binding motif, and that modification of a serine residue near a YEKV tyrosine did not eliminate interaction with β-catenin. From this, the examiner finds that "Li neither teaches away, discredits or otherwise discourage[s] the ordinary artisan from determining the role tyrosine phosphorylation may play in the interaction between MUC-1 and β-catenin." This very statement highlights the improper nature of the rejection. The claimed invention is not a method of determining ***whether*** tyrosines generally play a role, but assessing the effects of compounds on this action ***after*** it was determined that a specific tyrosine ***does*** play a role.

Third, the examiner makes a similar misapplication of the teachings of Yamamoto. As acknowledged, Yamamoto acknowledged that "it is not known if tyrosine sites influence binding of catenins to the serine rich motif." A more equivocal statement can hardly be imagined. Yet somehow, the examiner contorts this quote to into a "suggest[ion that] the phosphorylation of one or more of the seven tyrosine residues in the MUC1 cytoplasmic domain ... [is a] possible regulatory feature, wherein the YEKV site is immediately adjacent to the serine rich motif." To call this statement rank speculation would be too kind – it is nothing short of an outright misrepresentation of the teachings of the reference, as the previous quote from Yamamoto clearly disavows any evidence that tyrosines, much less YEKV tyrosines, are involved. The examiner, knowing this, hedges his bet by stating that "those of ordinary skill in the art were motivated to determine if other phosphorylated residues in the MUC1 cytoplasmic domain were responsible for the interaction between MUC1 and β-catenin." Again, appellants are ***not*** claiming to "determine" whether phosphorylated MUC1 residues have an impact on function, which this language would imply. Instead, they are claiming to exploit the finding, made by the inventors and ***not*** by Li or Yamamoto, that YEKV ***is*** in fact critical to MUC1's interaction with β-catenin. Without this knowledge, the prior art at best the art leaves one to pursue a general line of research that may or may not lead to fruition. This does not qualify as obvious. *In re O'Farrell*, 853 F.2d 894, 903, 7 USPQ2d 1673, 1681 (Fed. Cir. 1988).

Fourth, the examiner argues that appellants have overlooked the "emphasis" Zrihan "postulated" that MUC1 tyrosines interact with SH2 domain-containing proteins, while admitting that YEEV motifs are preferred. Thus, the examiner argues that "it does not teach away from all other tyrosines." Whether or not this is true, it highlights the fact that Zrihan certainly does not ***suggest*** the significance of YEKV motifs, and ***that*** is what is being claimed

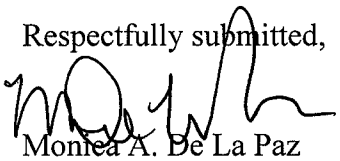
here. Thus, this reference too lacks any reasonable teaching or inference that would guide the skilled artisan to YEKV. At most, this is an invitation to invent, and it certainly cannot obviate appellants' invention based on that alone.

In conclusion, appellants submit that the following summation, offered by the examiner, highlights the baseless nature of the rejection: "The tyrosine phosphorylation of MUC1, and the YEKV site in particular, ***necessarily flows*** from the signal transduction pathways in cancer cells of Li *et al.* (1998), Yamamoto *et al.* and Barker." This language smacks of a inherency theory, which has no basis in an obviousness rejection. The examiner is simply grasping at straws in an vain effort to support a rejection that lacks the required teaching, suggestion and motivation in the cited art. In the end, one of ordinary skill would have no reasonable expectation of success that phosphorylation of a YEKV site would be important for interaction with β-catenin. Therefore, the Examiner has not set forth a *prima facie* case of obviousness.

C. Conclusion

In light of the foregoing, appellants respectfully submit that all pending claims are non-obvious under 35 U.S.C. §103. Therefore, it is respectfully requested that the Board reverse the pending rejection.

Respectfully submitted,



Monica A. De La Paz
Reg. No. 54,662
Attorney for Applicants

FULBRIGHT & JAWORSKI L.L.P.
600 Congress Avenue, Suite 2400
Austin, Texas 78701
(512) 474-5201
(512) 536-4598 (facsimile)

Date: August 25, 2009

VIII. APPENDIX A – APPEALED CLAIMS

1. A method of identifying a compound that inhibits binding of MUC1 to a tumor progressor, the method comprising:

(a) providing a MUC1 test agent, wherein the MUC1 test agent comprises a phosphorylated YEKV site;

(b) providing a tumor progressor test agent that binds to the phosphorylated MUC1 test agent;

(c) contacting the phosphorylated MUC1 test agent with the tumor progressor test agent in the presence of a test compound; and

(d) determining whether the test compound inhibits binding of the phosphorylated MUC1 test agent to the tumor progressor test agent.

5. The method of claim 1, wherein the tumor progressor test agent is a β-catenin test agent.

7. The method of claim 1, wherein the contacting is carried out in a cell-free system.

8. The method of claim 1, wherein the contacting occurs in a cell.

9. The method of claim 1, wherein the test compound is a peptide fragment of the tumor progressor.

13. The method of claim 9, wherein the tumor progressor test agent is a β-catenin test agent.

15. The method of claim 9, wherein the contacting is carried out in a cell-free system.

16. The method of claim 9, wherein the contacting occurs in a cell.

17. The method of claim 1, wherein the MUC1 test agent comprises SEQ ID NO:1.

22. The method of claim 8, wherein the cell is a cancer cell.

23. The method of claim 22, wherein the cancer cell expresses MUC1.

24. The method of claim 22, wherein the cancer cell is a breast cancer cell, a lung cancer cell, a colon cancer cell, a pancreatic cancer cell, a renal cancer cell, a stomach cancer cell, a liver cancer cell, a bone cancer cell, a hematological cancer cell, a neural tissue cancer cell, a melanoma cell, an ovarian cancer cell, a testicular cancer cell, a prostate cancer cell, a cervical cancer cell, a vaginal cancer cell, or a bladder cancer cell.

25. The method of claim 5, wherein providing a phosphorylated MUC1 test agent comprises combining a MUC1 test agent, a tumor progressor test agent with kinase activity, and ATP, wherein a MUC1 test agent phosphorylated at a YEKV site is formed.

26. The method of claim 25, wherein the tumor progressor test agent with kinase activity is c-src, EGF-R, or PKCδ.

IX. APPENDIX B – EVIDENCE CITED

Exhibit 1 – Li *et al.* (1998)

Exhibit 2 – Yamamoto *et al.* (1997)

Exhibit 3 – Barker *et al.* (U.S. Patent 5,851,775)

Exhibit 4 – Zrihan-Licht *et al.* (1994)

X. APPENDIX C – RELATED PROCEEDINGS

None

Accelerated Publication

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 276, No. 38, Issue of September 21, pp. 35239–35242, 2001

The Epidermal Growth Factor Receptor Regulates Interaction of the Human DF3/MUC1 Carcinoma Antigen with c-Src and β-Catenin*

Received for publication, June 26, 2001,and in revised form, July 30, 2001
Published, JBC Papers in Press, August 1, 2001,
DOI 10.1074/jbc.C100359200

Yongqing Li‡, Jian Ren‡, Wei-hsuan Yu§, Quan Li, Hiroaki Kuwahara, Li Yin, Kermit L. Carraway III¶, and Donald Kufe‖

From the Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts 02115, the §Division of New Born Medicine, Children's Hospital, Harvard Medical School, Boston, Massachusetts 02115, and the ¶Department of Biological Chemistry and University of California Davis Cancer Center, University of California at Davis, Sacramento, California 95817

The DF3/MUC1 mucin-like, transmembrane glycoprotein is aberrantly overexpressed in most human carcinomas. The MUC1 cytoplasmic domain interacts with the c-Src tyrosine kinase and thereby increases binding of MUC1 and β-catenin. In the present work, coimmunoprecipitation studies demonstrate that MUC1 associates constitutively with the epidermal growth factor receptor (EGF-R) in human ZR-75-1 breast carcinoma cells. Immunofluorescence studies show that EGF-R and MUC1 associate at the cell membrane. We also show that the activated EGF-R phosphorylates the MUC1 cytoplasmic tail on tyrosine at a YEKV motif that functions as a binding site for the c-Src SH2 domain. The results demonstrate that EGF-R-mediated phosphorylation of MUC1 induces binding of MUC1 to c-Src in cells. Moreover, *in vitro* and *in vivo* studies demonstrate that EGF-R increases binding of MUC1 and β-catenin. These findings support a novel role for EGF-R in regulating interactions of MUC1 with c-Src and β-catenin.

The epidermal growth factor receptor (EGF-R, HER1)[1] is a member of a family of transmembrane receptor tyrosine kinases that includes HER2/neu, HER3, and HER4 (1). EGF-R is activated by EGF, transforming growth factor-α, amphiregulin, and betacellulin. Following ligand binding, inactive monomeric EGF-R undergoes homodimerization or heterodimerization with other members of the HER family (2). Activation of EGF-R is associated with phosphorylation of specific tyrosine residues in the cytoplasmic region and thereby the recruitment of effector proteins that contain SH2 domains. For example, interaction of EGF-R with the Shc and Grb2 adaptor proteins links receptor activation to the Ras signaling pathway (3–5). Activation of EGF-R is also associated with the formation of complexes with the c-Src nonreceptor tyrosine kinase (6, 7). The finding that overexpression of EGF-R in fibroblasts confers growth in soft agar and induces tumorigenicity in nude mice has indicated that EGF-R can function as an oncogene (8, 9). Other studies in cells overexpressing both EGF-R and c-Src have shown that c-Src potentiates EGF-R-mediated tumorigenesis (7). The interaction between EGF-R and c-Src is further supported by the demonstration that c-Src is required for EGF-R-dependent mitogenesis (10).

The human DF3/MUC1 mucin-like glycoprotein is highly overexpressed by human carcinomas (11). Whereas MUC1 expression is restricted to the apical borders of normal secretory epithelium, this transmembrane protein is aberrantly expressed by carcinoma cells at high levels over the entire cell surface (11). The MUC1 protein consists of an N-terminal ectodomain with variable numbers of conserved 20 amino acid tandem repeats that are subject to *O*-glycosylation (12, 13). The C-terminal region includes a transmembrane domain and a 72-amino acid cytoplasmic tail that contains seven sites for tyrosine phosphorylation. The >250-kDa ectodomain associates with the ~25 kDa C-terminal region as a heterodimer at the cell surface. β-Catenin, a component of the adherens junction of mammalian epithelial cells, binds directly to MUC1 at a SAGNGGSSL motif in the cytoplasmic domain (14). Similar motifs are responsible for interactions of β-catenin with E-cadherin and the adenomatous polyposis coli tumor suppressor (15–17). Glycogen synthase kinase 3β (GSK3β) also binds to MUC1 and phosphorylates serine in a SPYEKV site adjacent to that for the β-catenin interaction (18). More recent studies have shown that c-Src phosphorylates the SPYEKV site on tyrosine (19). The findings also demonstrate that c-Src increases, while GSK3β down-regulates, the interaction between MUC1 and β-catenin (18, 19).

The present studies demonstrate that the EGF-R interacts with MUC1. The activated EGF-R phosphorylates MUC1 on the YEKV motif in the cytoplasmic tail. The results also demonstrate that EGF-R-mediated phosphorylation of MUC1 induces binding of MUC1 to c-Src and β-catenin.

MATERIALS AND METHODS

Cell Culture—Human ZR-75-1 carcinoma cells were grown in RPMI 1640 medium containing 10% heat-inactivated fetal bovine serum (HI-FBS), 100 units/ml penicillin, 100 μg/ml streptomycin, and 2 mM L-glutamine. HCT116 and 293 cells were cultured in Dulbecco's modified Eagle's medium with 10% HI-FBS, 100 units/ml penicillin and 100 μg/ml streptomycin. In certain studies, cells were cultured in medium with 0.1% HI-FBS for 24 h and then stimulated with 10 ng/ml EGF (Calbiochem-Novabiochem, San Diego, CA) for 5 min at 37 °C.

Cell Transfections—Wild-type MUC1 containing 40 tandem repeats was excised from pCMV-MUC1 (19) by *Nde*I and *Eco*RI digestion and integrated into the *NdeI/Eco*RI site of the mammalian expression vector pIRESpuro2 (CLONTECH, Palo Alto, CA). The pIRESpuro2-MUC1 (Y46F) mutant vector was constructed by insertion of the 3′-terminal region from pCMV-MUC1(Y46F) (19) into pIRESpuro2-MUC1 deleted at the 3′-terminal region of MUC1 by Bsu36I. 293 cells were transiently transfected with pcDNA3.1/EGF-R and/or pIRESpuro2-MUC1 by Lipo-

* This work was supported by Grant CA87421 awarded by the NCI. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ These authors contributed equally to this work.
‖ To whom correspondence should be addressed: Adult Oncology, Dana-Farber Cancer Institute, 44 Binney St., Boston, MA 02115. Tel.: 617-632-3141; Fax: 617-632-2934; E-mail: donald_kufe@dfci.harvard.edu.

[1] The abbreviations used are: EGF, epidermal growth factor; EGF-R, EGF receptor; GSK3β, glycogen synthase 3β; CD, cytoplasmic domain; HI-FBS, heat-inactivated fetal bovine serum; PAGE, polyacrylamide gel electrophoresis; GST, glutathione *S*-transferase.

fectAMINE (Life Technologies Inc.). Cell lysates were prepared at 48 h after transfection. pIRESpuro2, pIRESpuro2-MUC1, and pIRESpuro2-MUC1(Y46F) were transfected into HCT116 colon carcinoma cells by LipofectAMINE. Stable transfectants were selected in the presence of 0.4 μg/ml of puromycin (Calbiochem-Novabiochem Co., San Diego, CA).

Lysate Preparation—Lysates from subconfluent cells were prepared as described previously (18).

Immunoprecipitation and Immunoblotting—Equal amounts of protein from the cell lysates were incubated with mouse or rabbit IgG, monoclonal antibody DF3 (anti-MUC1) (11), or anti-EGF-R (Santa Cruz Biotechnology, Santa Cruz, CA). The immune complexes were prepared as described previously (18), separated by SDS-PAGE, and transferred to nitrocellulose membranes. The immunoblots were probed with monoclonal antibody DF3, anti-EGF-R, anti-c-Src (Upstate Biotechnology, Lake Placid, NY), anti-P-Tyr (RC20H; Transduction Laboratories, San Diego, CA), or anti-β-catenin (Zymed Laboratories Inc., San Francisco, CA). Reactivity was detected with horseradish peroxidase-conjugated second antibodies and chemiluminescence (PerkinElmer Life Sciences).

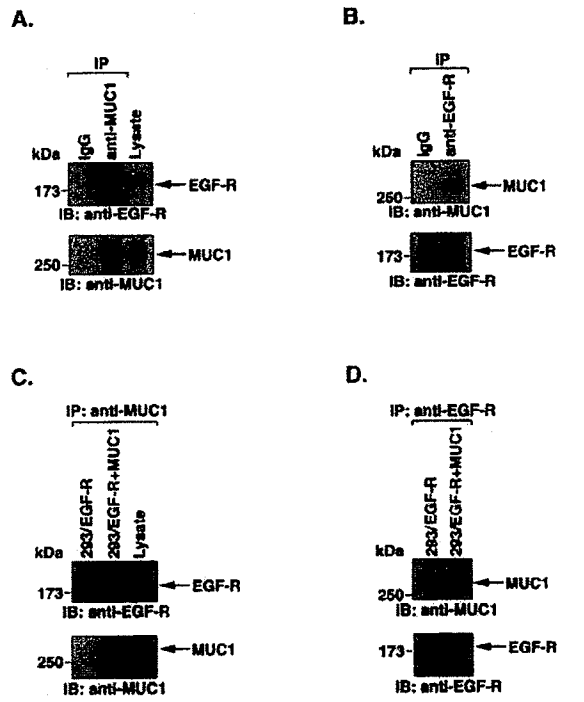


FIG. 1. **Interaction of MUC1 with EGF-R.** *A* and *B*, lysates from ZR-75–1 cells were subjected to immunoprecipitation (*IP*) with anti-MUC1 (*A*) or anti-EGF-R (*B*). Rabbit or mouse IgG was used as a control. The immunoprecipitates were analyzed by immunoblotting (*IB*) with anti-EGF-R and anti-MUC1. *C* and *D*, 293 cells were transiently transfected with EGF-R or EGF-R + MUC1. At 48 h after transfection, the cells were harvested, and lysates were subjected to immunoprecipitation with anti-MUC1 (*C*) or anti-EGF-R (*D*). The immunoprecipitates were analyzed by immunoblotting with anti-EGF-R and anti-MUC1.

Immunofluorescence Microscopy—ZR-75-1 cells were fixed with 4% paraformaldehyde for 10 min at room temperature and blocked with 5% fatty acid-free BSA (Sigma) and 5% normal goat serum (Jackson Immunoresearch Laboratories Inc., Westgrove, PA) in phosphate-buffered saline (blocking buffer) for 45 min at room temperature. After incubation with anti-MUC1 (1:400) and rabbit anti-EGF-R (1:100) in blocking buffer for 14 h at 4 °C, the cells were washed with phosphate-buffered saline and incubated with fluorescein-conjugated anti-rabbit IgG (1:100) or Texas Red-conjugated anti-mouse IgG (1:200) (Jackson Immunoresearch Laboratories Inc., Westgrove, PA) for 45 min at room temperature. The cells were then mounted onto coverslips using the slow fade mounting kit (Molecular Probes, Eugene, OR) and analyzed by confocal microscopy (inverted Zeiss LSM 510). Images were captured at 0.6-nm increments along the *z* axis under ×63 magnification and converted to composite images by ImageSpace 3.10 software (Molecular Dynamics, Sunnyvale, CA).

In Vitro Phosphorylation—Purified wild-type and mutant MUC1/CD proteins were incubated with 0.1 unit of purified EGF-R (Calbiochem-Novabiochem Co, San Diego, CA) in 20 μl of kinase buffer (20 mM Tris-HCl, pH 7.6, 10 mM $MgCl_2$, 5 mM dithiothreitol). Kinase reactions and analysis of the reaction products were performed as described previously (19).

Binding Studies—Purified wild-type and mutant MUC1/CD proteins were incubated with 0.1 unit of EGF-R in the absence and presence of 200 μM ATP for 30 min at 30 °C. GST-c-Src or GST-β-catenin bound to glutathione beads was then added, and the reaction was incubated for 1 h at 4 °C. After washing, the precipitated proteins were subjected to immunoblot analysis with anti-MUC1/CD (18) or anti-Tyr(P).

RESULTS AND DISCUSSION

MUC1 Associates with EGF-R—To determine whether MUC1 forms a complex with EGF-R, anti-MUC1 immunoprecipitates from lysates of human ZR-75-1 cells were analyzed by immunoblotting with anti-EGF-R. The results demonstrate that EGF-R coprecipitates with MUC1 (Fig. 1*A*). As a control, there was no detectable EGF-R in immunoprecipitates prepared with IgG (Fig. 1*A*). In the reciprocal experiment, analysis of anti-EGF-R immunoprecipitates with anti-MUC1 confirmed that EGF-R associates with MUC1 (Fig. 1*B*). To extend these findings, 293 cells, which express low levels of EGF-R and are negative for MUC1 (18), were transfected to express EGF-R and MUC1. Immunoblot analysis of anti-MUC1 immunoprecipitates with anti-EGF-R demonstrated coprecipitation of

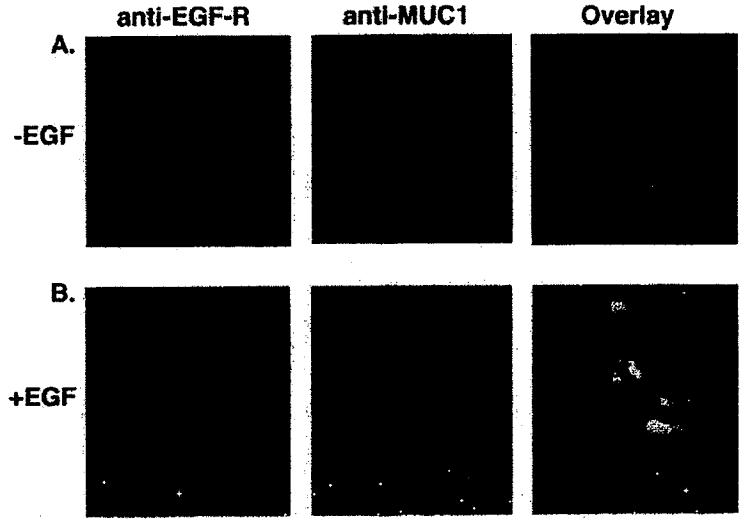


FIG. 2. **Colocalization of EGF-R and MUC1 in ZR-75-1 cells.** ZR-75-1 cells were grown to 80% confluence in a Lab-Tek II Chamber Slide and then incubated in medium with 0.1% serum for 24 h. The cells were treated without (*A*) or with (*B*) 10 ng/ml EGF for 5 min, fixed, and double-stained with anti-EGF-R (*green*) and anti-MUC1 (*red*).

EGF-R and MUC1 (Fig. 1C). Similar results were obtained when anti-EGF-R immunoprecipitates were analyzed by immunoblotting with anti-MUC1 (Fig. 1D). These findings demonstrate that MUC1 constitutively associates with EGF-R.

Colocalization of EGF-R and MUC1 by Immunofluorescence Microscopy—To assess the subcellular localization of MUC1 and EGF-R, confocal microscopy was performed with rabbit anti-EGF-R and mouse anti-MUC1 antibodies. In control ZR-75-1 cells, EGF-R was distributed uniformly over the cell membrane (Fig. 2A, *left*). Similar findings were obtained for the distribution of MUC1 (Fig. 2A, *middle*). Overlay of the EGF-R (*green*) and MUC1 (*red*) signals supported colocalization (*red* + *green* → *yellow*) (Fig. 2A, *right*). Following EGF stimulation, the EGF-R signals were clustered in patches at the cell membrane (Fig. 2B, *left*). An identical pattern was observed for MUC1 (Fig. 2B, *middle*). Moreover, overlay of the signals showed that EGF-R and MUC1 colocalize in clusters at the cell membrane (Fig. 2B, *right*). Analysis of the control and EGF-stimulated cells by coimmunoprecipitation studies demonstrated no detectable difference in the association between EGF-R and MUC1 (data not shown). These findings and those obtained in coprecipitation studies demonstrate that MUC1 and EGF-R associate constitutively at the cell membrane.

EGF-R Phosphorylates MUC1 in Vitro and in Vivo—To determine whether EGF-R phosphorylates MUC1, anti-MUC1 immunoprecipitates from control and EGF-stimulated ZR-75-1 cells were analyzed by immunoblotting with anti-Tyr(P). The results demonstrate a detectable level of tyrosine-phosphorylated MUC1 in control cells (Fig. 3A). Moreover, EGF stimulation was associated with an increase in phosphorylation of MUC1 on tyrosine (Fig. 3A). EGF-induced tyrosine phosphorylation of MUC1 was also observed in 293 cells transfected to express EGF-R and MUC1 (Fig. 3B). The 72-amino acid MUC1 cytoplasmic domain (MUC1/CD) contains 7 tyrosines (see schema in Fig. 4D). To define potential sites of EGF-R phosphorylation, we incubated the MUC1 cytoplasmic domain (MUC1/CD) with EGF-R and [γ-^{32}P]ATP. Analysis of the reaction products demonstrated that EGF-R phosphorylates MUC1/CD (Fig. 3C). Mutation of the Tyr^{8} site to Phe had no detectable effect on EGF-R-mediated phosphorylation of MUC1/CD (Fig. 3C). There was also no apparent effect when the Tyr^{20} or Tyr^{35} sites were mutated to Phe (Fig. 3C). By contrast, incubation of MUC1/CD(Y46F) with EGF-R was associated with a marked decrease in phosphorylation as compared with that found with wild-type MUC1/CD (Fig. 3C). Mutation of Tyr^{26} also resulted in decreased phosphorylation, but to a lesser extent than that obtained with Y46F (Fig. 3C). To determine whether the Tyr^{46} site is phosphorylated *in vivo*, human HCT116 cells, which express EGF-R and not MUC1, were stably transfected to express the empty vector, wild-type MUC1, or the MUC1(Y46F) mutant. Analysis of anti-MUC1 immunoprecipitates with anti-Tyr(P) demonstrated that EGF-mediated phosphorylation of MUC1(Y46F) is decreased compared with that obtained with wild-type MUC1 (Fig. 3D). Relative intensities of the anti-Tyr(P) signals were determined by densitometric scanning (Fig. 3D). Similar results were obtained in three separate experiments (legend to Fig. 3D). The findings that the MUC1(Y46F) mutation decreases tyrosine phosphorylation only in part is in concert with additional tyrosine sites in the MUC1/CD, which can function as substrates for other tyrosine kinases. These results thus demonstrate that EGF-R phosphorylates MUC1 on Tyr^{46} *in vitro* and in cells.

EGF-R Regulates Interaction of MUC1 with c-Src and β-Catenin—To determine whether EGF-R-mediated phosphorylation regulates the interaction of MUC1 with c-Src and β-catenin, we incubated MUC1/CD with EGF-R and ATP and

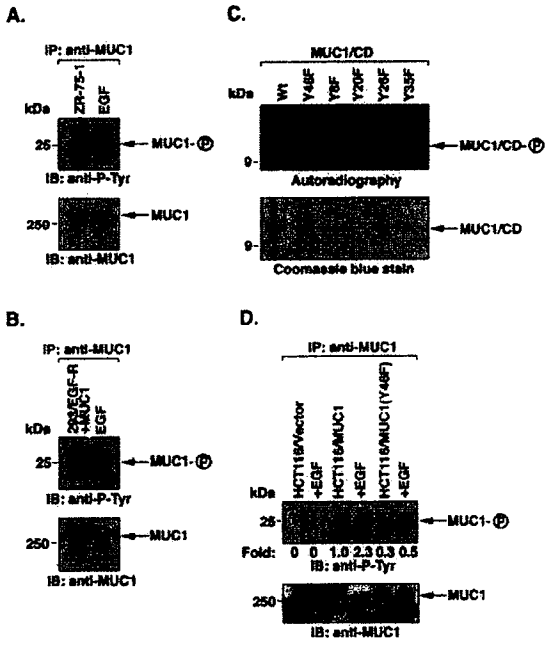


FIG. 3. **EGF-R phosphorylates MUC1 *in vitro* and *in vivo*.** *A*, lysates from ZR-75-1 cells treated with or without 10 ng/ml EGF for 5 min were subjected to immunoprecipitation with anti-MUC1. The immunoprecipitates were analyzed with anti-Tyr(P) (*upper panel*) and anti-MUC1 (*lower panel*). *B*, 293 cells were transiently transfected with EGF-R and MUC1. At 48 h, the cells were treated with or without 10 ng/ml EGF for 5 min. Lysates were subjected to immunoprecipitation with anti-MUC1. The immunoprecipitates were analyzed by immunoblotting with anti-Tyr(P) and anti-MUC1. *C*, purified wild-type MUC1/CD and the indicated mutant proteins were incubated with purified EGF-R and [γ-^{32}P]ATP. The reaction products were analyzed by SDS-PAGE and autoradiography (*upper panel*). Equal loading of the proteins was assessed by Coomassie Blue staining (*lower panel*). *D*, lysates from HCT116/vector, HCT116/MUC1, and HCT116/MUC1(Y46F) cells treated without or with 10 ng/ml EGF for 5 min at 37 °C were subjected to immunoprecipitation with anti-MUC1. The immunoprecipitates were analyzed by immunoblotting with anti-Tyr(P) and anti-MUC1. The signal intensities of the tyrosine-phosphorylated MUC1 proteins were compared with that of control HCT116/MUC1 (designated 1.0) by densitometric scanning. The relative intensities (mean ± S.E.) from three separate experiments were: HCT116/vector, 0 ± 0; HCT116/vector +EGF, 0 ± 0; HCT116/MUC1, 1.0; HCT116/MUC1 +EGF, 2.2 ± 0.3; HCT116/MUC1(Y46F), 0.5 ± 0.3; and HCT116/MUC1(Y46F), 0.6 ± 0.1.

then assessed binding to GST-Src-SH2 and GST-β-catenin. Immunoblot analysis of adsorbates to glutathione beads with anti-MUC1/CD showed that GST-Src SH2 binds to MUC1/CD following EGF-R phosphorylation (Fig. 4A). In addition, compared with MUC1/CD, there was substantially less binding of GST-Src-SH2 to the MUC1/CD(Y46F) mutant that had been incubated with EGF-R and ATP (Fig. 4A). Similar findings were obtained for binding of GST-β-catenin (Fig. 4A). To assess whether EGF-R-mediated phosphorylation of MUC1 induces binding of MUC1 to c-Src and β-catenin *in vivo*, anti-MUC1 immunoprecipitates from ZR-75-1 cells were analyzed by immunoblotting with anti-c-Src or anti-β-catenin. Analysis of lysates from control ZR-75-1 cells demonstrated a low but detectable interaction of MUC1 with c-Src and β-catenin (Fig. 4B). In concert with the *in vitro* results, stimulation of ZR-75-1 cells with EGF induced the interaction of MUC1 with c-Src and β-catenin (Fig. 4B). To confirm involvement of the MUC1 Tyr^{46} site, HCT116 cells stably expressing wild-type MUC1 or MUC1(Y46F) were stimulated with EGF. Immunoblot analysis of anti-MUC1 immunoprecipitates with anti-c-Src demonstrated that, compared with wild-type MUC1, there was less EGF-induced binding

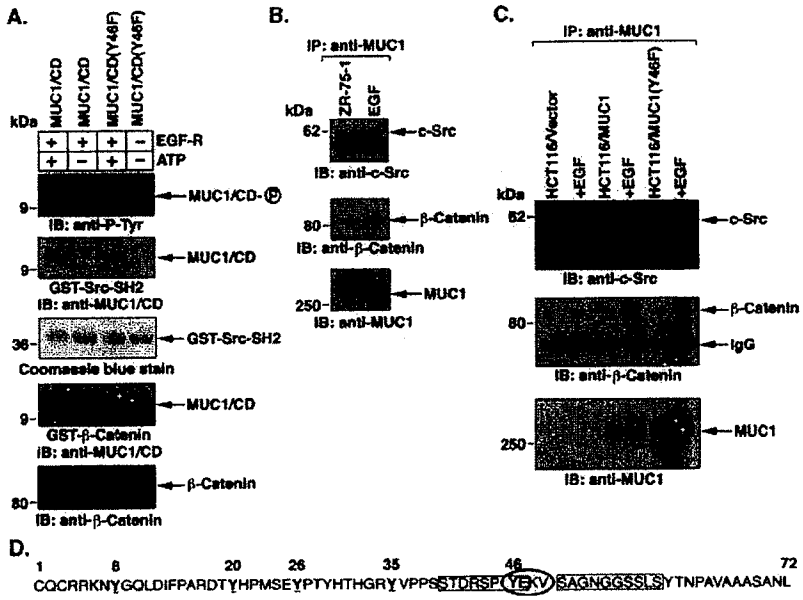


FIG. 4. **EGF-R-mediated phosphorylation of MUC1 enhances the interaction of MUC1 with c-Src and β-catenin.** *A*, purified MUC1/CD or MUC1/CD(Y46F) was incubated with EGF-R in the presence or absence of ATP for 30 min at 30 °C. GST-Src-SH2 or GST-β-catenin was then added, and the reaction was incubated for 1 h at 4 °C. Proteins precipitated with glutathione beads were separated by SDS-PAGE and subjected to immunoblot analysis with anti-MUC1/CD and anti-Tyr(P). Equal loading of the proteins was assessed by Coomassie Blue staining or immunoblotting with anti-β-catenin. *B*, lysates from ZR-75-1 cells treated with or without 10 ng/ml EGF for 5 min were subjected to immunoprecipitation with anti-MUC1. The immunoprecipitates were analyzed by immunoblotting with anti-c-Src (*upper panel*), anti-β-catenin (*middle panel*), and anti-MUC1 (*lower panel*). *C*, HCT116/vector, HCT116/MUC1, and HCT116/MUC1(Y46F) cells were treated without or with 10 ng/ml EGF for 5 min at 37 °C. Lysates were subjected to immunoprecipitation with anti-MUC1. The immunoprecipitates were analyzed by immunoblotting with anti-c-Src (*upper panel*), anti-β-catenin (*middle panel*), and anti-MUC1 (lower panel). *D*, amino acid sequence of the MUC1 cytoplasmic tail. The GSK3β binding and phosphorylation site (STDRSP), the c-Src SH2 binding motif (YEKV), and the β-catenin binding site (SAGNGGSSLS) are highlighted.

of MUC1(Y46F) to c-Src (Fig. 4*C*). Similar findings were obtained for β-catenin (Fig. 4*C*). These results show that EGF-R-mediated phosphorylation of MUC1 Y46 induces the interaction of MUC1 with c-Src and β-catenin.

MUC1 Integrates EGF-R, c-Src, and β-Catenin Signaling—The present findings and those recently reported for the mouse mammary gland (20) demonstrate that MUC1 interacts with EGF-R. The present results further show that EGF-R phosphorylates the MUC1 cytoplasmic tail on the YEKV motif. These findings and the recent demonstration that c-Src phosphorylates Tyr^{46} (19) have supported regulation of the YEKV site by both EGF-R and c-Src. The available evidence indicates that Tyr^{46} is functionally important in regulating the interactions of MUC1 with multiple signaling pathways. EGF-R-mediated phosphorylation of MUC1 Tyr^{46} functions as a binding site for the c-Src SH2 domain. Moreover, EGF stimulation is associated with increased binding of MUC1 and c-Src *in vivo*. EGF-R-mediated phosphorylation of MUC1 Tyr^{46} also induces binding of MUC1 with β-catenin *in vitro* and in cells. In addition, phosphorylation of MUC1 Tyr^{46} by EGF-R or c-Src down-regulates the interaction of MUC1 and GSK3β (Ref. 19 and data not shown). In concert with these findings, the YEKV motif resides between the GSK3β binding and phosphorylation site (STDRSP) (18) and the β-catenin binding site (SAGNGGSSLS) (14) (Fig. 4*D*). Taken together, the present results and those of previous studies (14, 18–20) support a model in which EGF stimulation induces phosphorylation of MUC1 on Tyr^{46} and thereby integrates signaling among the c-Src, β-catenin, and GSK3β pathways. The aberrant overexpression of MUC1 in human carcinoma cells could thus contribute to the transformed phenotype by dysregulation of EGF-R, c-Src, β-catenin, and/or GSK3β signaling.

Acknowledgment—We are grateful to Kamal Chauhan for his excellent technical support.

REFERENCES

1. van der Geer, P., Hunter, T., and Lindberg, R. A. (1994) *Annu. Rev. Cell Biol.* **10**, 251–337
2. Yarden, Y., and Schlessinger, J. (1987) *Biochemistry* **26**, 1443–1451
3. Li, N., Batzer, A., Daly, R., Yajnik, V., Skolnik, E., Chardin, P., Bar-Sagi, D., Margolis, B., and Schlessinger, J. (1993) *Nature* **363**, 85–88
4. Egan, S. E., Giddings, B. W., Brooks, M. W., Buday, L., Sizeland, A. M., and Weinberg, R. A. (1993) *Nature* **363**, 45–51
5. Rozakis-Adcock, M., McGlade, J., Mbamalu, G., Pelicci, G., Daly, R., Li, W., Batzer, A., Thomas, S., Brugge, J., Pelicci, P. G., Schlessinger, J., and Pawson, T. (1992) *Nature* **360**, 689–692
6. Luttrell, D. K., Lee, A., Lansing, T. J., Crosby, R. M., Jung, K. D., Willard, D., Luther, M., Rodriguez, M., Berman, J., and Gilmer, T. M. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91**, 83–87
7. Maa, M., Leu, T., McCarley, D., Schatzman, R., and Parsons, S. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92**, 6981–6985
8. Di Fiore, P. P., Pierce, J. H., Fleming, T. P., Hazan, R., Ullrich, A., King, C. R., Schlessinger, J., and Aaronson, S. A. (1987) *Cell* **51**, 1063–1070
9. Velu, T. J., Beguinot, L., Vass, W. C., Willingham, M. C., Merlino, G. T., Pastan, I., and Lowy, D. R. (1987) *Science* **238**, 1408–1410
10. Wilson, L. K., Luttrell, D. K., Parsons, J. T., and Parsons, S. J. (1989) *Mol. Cell. Biol.* **9**, 1536–1544
11. Kufe, D., Inghirami, G., Abe, M., Hayes, D., Justi-Wheeler, H., and Schlom, J. (1984) *Hybridoma* **3**, 223–232
12. Gendler, S., Taylor-Papadimitriou, J., Duhig, T., Rothbard, J., and Burchell, J. A. (1988) *J. Biol. Chem.* **263**, 12820–12823
13. Siddiqui, J., Abe, M., Hayes, D., Shani, E., Yunis, E., and Kufe, D. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 2320–2323
14. Yamamoto, M., Bharti, A., Li, Y., and Kufe, D. (1997) *J. Biol. Chem.* **272**, 12492–12494
15. Rubinfeld, B., Souza, B., Albert, I., Muller, O., Chamberlain, S. H., Masiarz, F. R., Munemitsu, S., and Polakis, P. (1993) *Science* **262**, 1731–1734
16. Su, L.-K., Vogelstein, B., and Kinzler, K. W. (1993) *Science* **262**, 1734–1737
17. Rubinfeld, B., Souza, B., Albert, I., Munemitsu, S., and Polakis, P. (1995) *J. Biol. Chem.* **270**, 5549–5555
18. Li, Y., Bharti, A., Chen, D., Gong, J., and Kufe, D. (1998) *Mol. Cell. Biol.* **18**, 7216–7224
19. Li, Y., Kuwahara, H., Ren, J., Wen, G., and Kufe, D. (2001) *J. Biol. Chem.* **276**, 6061–6064
20. Schroeder, J., Thompson, M., Gardner, M., and Gendler, S. (2001) *J. Biol. Chem.* **276**, 13057–13064

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 272, No. 19, Issue of May 9, pp. 12492–12494, 1997
Printed in U.S.A.

Interaction of the DF3/MUC1 Breast Carcinoma-associated Antigen and β-Catenin in Cell Adhesion*

(Received for publication, March 5, 1997)

Makiko Yamamoto, Ajit Bharti, Yongqing Li, and Donald Kufe

From the Division of Cancer Pharmacology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts 02115

The DF3/MUC1 mucin-like glycoprotein is aberrantly overexpressed in human breast carcinomas. The functional role of DF3 is unknown. The present studies demonstrate that DF3 associates with β-catenin. Similar findings have been obtained for γ-catenin but not α-catenin. DF3, like E-cadherin and the adenomatous polyposis coli gene product, contains an SXXXXXSSL site that is responsible for direct binding to β-catenin. The results further demonstrate that interaction of DF3 and β-catenin is dependent on cell adhesion. These findings and the role of β-catenin in cell signaling support a role for DF3 in the adhesion of epithelial cells.

The human *DF3* (*MUC1*, *episialin*, *PEM*) gene encodes a high molecular mass membrane-associated glycoprotein with a mucin-like external domain. The DF3 glycoprotein is expressed on the apical borders of secretory mammary epithelial cells and aberrantly expressed over the entire surface of carcinoma cells (1). The ectodomain consists of varying numbers of 20-amino acid tandem repeats that are subject to *O*-glycosylation and that contribute to the expression of a polymorphic protein (2–4). The N-terminal region contains hydrophobic signal sequences that vary as a consequence of alternate splicing (5–7). The C-terminal region includes a transmembrane domain and a 72-amino acid cytoplasmic tail that contains tyrosine phosphorylation sites (8, 9). The function of DF3 is unclear. However, high levels found on carcinoma cells reduce cell-cell and cell-extracellular matrix adhesion in a nonspecific manner (10–12). These studies have suggested that DF3 interferes with cellular adhesion by steric hindrance from the rigid ectodomain (11).

Cadherin cell adhesion molecules form complexes with the cytoplasmic α-, β-, and γ-catenin proteins (13). α-Catenin is required for cadherin-mediated cell adhesion and links cadherins to the actin cytoskeleton (14, 15). β-Catenin links α-catenin to the cadherins and is highly related to plakoglobin (γ-catenin) (16–18). β-Catenin is homologous to the *Drosophila* segment polarity gene product Armadillo (19) that acts downstream of Wingless (20). Armadillo forms complexes with *Drosophila* E-cadherin and α-catenin (21, 22). These findings have supported a role for β-catenin in morphogenetic signals. Other studies have demonstrated that β-catenin binds directly to the adenomatous polyposis coli (APC)[1] gene product (23–25). The APC protein and E-cadherin form independent complexes with β-catenin (25). γ-Catenin mediates similar interactions among APC, α-catenin, and the cytoskeleton (16).

The present results demonstrate that DF3 interacts directly with β-catenin. An SXXXXXSSL motif in the DF3 cytoplasmic domain is responsible for binding to β-catenin. We also demonstrate that cell adhesion induces the interaction between DF3 and β-catenin.

MATERIALS AND METHODS

Cell Culture—Human ZR-75-1 breast carcinoma cells were grown in RPMI 1640 medium containing 10% heat-inactivated fetal bovine serum, 100 μg/ml streptomycin, 100 units/ml penicillin, and 2 mM L-glutamine. Cells were grown in suspension (0.3×10^6/100 ml) with gentle rocking or as a monolayer on polystyrene culture dishes.

Cell Lysate—Cells (~70% confluent) were lysed in ice-cold lysis buffer (150 mM NaCl, 50 mM Tris, pH 7.6, 0.5% Brij 97, 10 μg/ml leupeptin, 10 μg/ml aprotinin, 1 mM sodium vanadate, 1 mM phenylmethylsulfonyl fluoride, and 1 mM dithiothreitol) for 30 min on ice. Lysates were cleared by centrifugation at $14,000 \times g$ for 15 min.

Immunoprecipitation and Immunoblotting—Lysates were incubated with monoclonal antibody (mAb) DF3 (1), anti-α-catenin (Zymed Laboratories, Inc., San Francisco, CA), anti-β-catenin (Zymed), anti-γ-catenin (Zymed), or anti-E-cadherin (Transduction Laboratories, Lexington, KY) for 2 h at 4 °C. Immunoprecipitates were prepared by incubation with rabbit anti-mouse IgG (Upstate Biotechnology, Inc., Lake Placid, NY) and protein A-Sepharose (Pharmacia Biotech Inc.) for 1 h at 4 °C. The precipitates were subjected to electrophoresis in 7.5% or 6% SDS-polyacrylamide gels. Proteins were transferred to nitrocellulose membranes by dry transfer. The membranes were blocked in 5% nonfat dry milk in phosphate-buffered saline containing 0.05% Tween 20 and then incubated with an appropriate antibody for immunoblot analysis. Reactivity was detected by horseradish peroxidase-conjugated second antibodies and chemiluminescence (ECL, Amersham Corp.).

Direct Binding Studies—The GST fusion construct expressing the DF3 cytoplasmic domain (CD) was prepared by polymerase chain reaction cloning and ligation into the pGEX2T vector. GST or GST-DF3/CD was affinity-purified with glutathione-Sepharose 4B beads and suspended in elution buffer (50 mM Tris-HCl, pH 8.0, 5 mM glutathione). Nitrocellulose filters were incubated with GST or GST-DF3/CD for 1.5 h at room temperature. Reactivity was detected with an anti-GST antibody (Santa Cruz Biotechnology).

RESULTS AND DISCUSSION

To identify proteins that associate with DF3, we analyzed mAb DF3 immunoprecipitates by SDS-PAGE and silver staining. The detection of a coprecipitated protein of 92 kDa was confirmed by reactivity with an antibody against β-catenin (Fig. 1*A*). Since E-cadherin forms complexes with α-, β-, and γ-catenins (26), we analyzed anti-DF3 immunoprecipitates for an association with α- and γ-catenins. While there was no detectable α-catenin in the precipitates, the results indicate that DF3 forms complexes with γ-catenin (Fig. 1, *B* and *C*). In the reciprocal experiments, anti-catenin immunoprecipitates were analyzed by immunoblotting with anti-DF3. The findings confirm binding of DF3 to β- and γ-catenins (Fig. 1*D*). As previously shown (26), E-cadherin formed complexes with all three of the catenins (Fig. 1*D*).

To determine if binding to DF3 is direct, we subjected anti-β-

* This work was supported by Department of the Army Grant DAMD 17-94-J-4394. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] The abbreviations used are: APC, adenomatous polyposis coli; mAb, monoclonal antibody; CD, cytoplasmic domain; GST, glutathione *S*-transferase; PAGE, polyacrylamide gel electrophoresis.

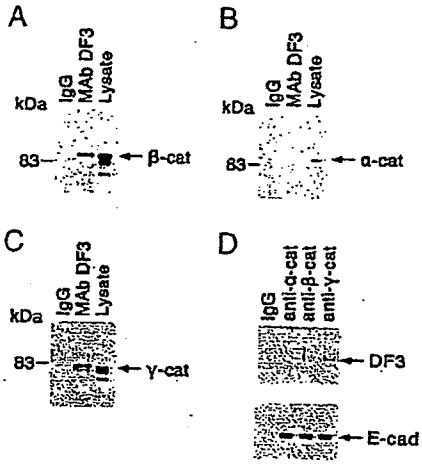


FIG. 1. Association of DF3 with β-catenin (*β-cat*) and γ-catenin. Lysates from adherent ZR-75–1 cells were subjected to immunoprecipitation with mAb DF3. The immunoprecipitates were analyzed for reactivity with anti-β-catenin (*A*), anti-α-catenin (*B*), and anti-γ-catenin (*C*). Lysates were directly analyzed by immunoblotting as controls. *D*, lysates were subjected to immunoprecipitation with the indicated antibodies. The precipitates were analyzed by immunoblotting with mAb DF3 (*upper panel*) or anti-E-cadherin (*E-cad, lower panel*).

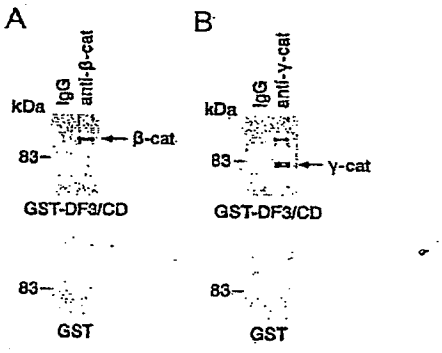


FIG. 2. Direct binding of DF3 to β-catenin (*β-cat*) and γ-catenin. Lysates were subjected to immunoprecipitation with anti-β-catenin (*A*) or anti-γ-catenin (*B*). The immunoprecipitates were separated by SDS-PAGE, and the proteins were transferred to nitrocellulose filters. The filters were incubated with GST or GST-DF3/CD and then washed and analyzed for reactivity with anti-GST.

catenin immunoprecipitates to SDS-PAGE and then transferred the separated proteins to filters. Incubation of the filters with a GST fusion protein that contains the DF3 cytoplasmic domain (GST-DF3/CD) demonstrated binding to β-catenin (Fig. 2*A*). By contrast, there was no detectable binding to GST (Fig. 2*A*). Similar results were obtained for γ-catenin (Fig. 2*B*).

Previous studies have demonstrated that β-catenin binds to SXXXXXSSL sites in E-cadherin (amino acids 840–848) and APC (seven motifs) (23, 24, 27) (Fig. 3*A*). β-Catenin also associates with the epidermal growth factor receptor, which contains a SRTPLLSSLS (amino acids 1030–1039) site (28). A similar site is present at amino acids 1239–1243 in DF3 (Fig. 3*A*). To assess whether β-catenin binds to the SXXXXXSSL site in DF3, we subjected cell lysates to immunoprecipitation with mAb DF3 in the presence of the synthetic peptide GGSSLSY. The results demonstrate that the peptide inhibits binding of β-catenin and DF3 (Fig. 3*B*). By contrast, there was no detectable effect on this interaction when using an irrelevant peptide (Fig. 3*B*). The GGSSLSY peptide also blocked interaction of DF3 and γ-catenin (Fig. 3*B*). These findings suggested that β- and γ-catenin bind to DF3 at similar sites.

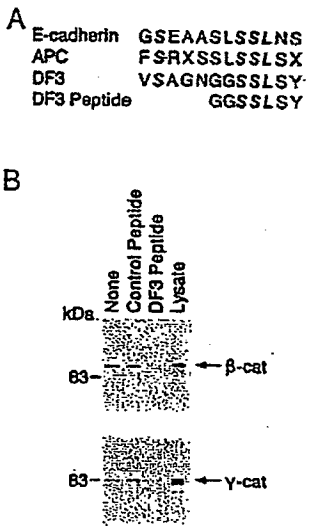


FIG. 3. DF3 binds to catenins at an SXXXXXSSL site. *A*, SXXXXXSSL sites in E-cadherin, APC, and DF3. *B*, lysates were subjected to immunoprecipitation with mAb DF3 in the presence of no added peptide, a control peptide (50 μM; EAPPPKIPDKQ), or a 50 μM GGSSLSY peptide. The immunoprecipitates were analyzed by immunoblotting with anti-β-catenin (*β-cat*) or anti-γ-catenin.

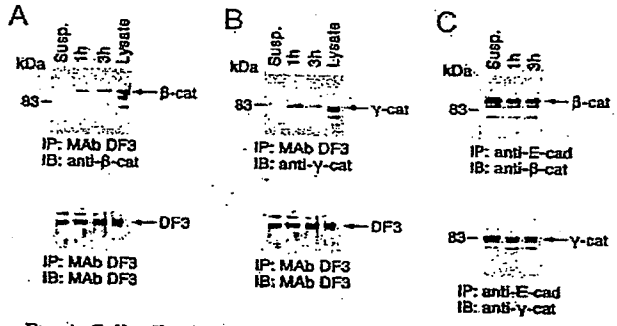


FIG. 4. Cell adhesion induces binding of DF3 with β-catenin (*β-cat*) and γ-catenin. Cells were trypsinized and grown in suspension by gentle agitation for 6 h. Suspension cells were allowed to adhere to culture dishes for 1 and 3 h. Lysates were subjected to immunoprecipitation with mAb DF3 (*A* and *B*) or anti-E-cadherin (*anti-E-cad, C*). The immunoprecipitates were analyzed for reactivity with anti-β-catenin, anti-γ-catenin, or mAb DF3. Lysates were directly analyzed by immunoblotting as controls.

The functional role of the association between DF3 and β-catenin was studied in cells grown in suspension and then grown as a monolayer. There was no detectable β-catenin in the mAb DF3 immunoprecipitates prepared from the suspension cells. By contrast, binding of DF3 to β-catenin was detectable at 1 and 3 h of adherence (Fig. 4*A*). Cell adhesion was also associated with formation of a complex with DF3 and γ-catenin (Fig. 4*B*), but not α-catenin (data not shown). A similar analysis of E-cadherin immunoprecipitates demonstrated little if any difference in binding to β- or γ-catenin in suspension as compared with adherent cells (Fig. 4*C*).

β-Catenin is involved in the formation of adherens junctions of epithelial cells. The cell adhesion E-cadherin protein and the APC tumor suppressor gene product compete for binding to the arm repeats of β-catenin (16) that are also found in Armadillo, γ-catenin, and certain other junctional proteins (29). The present studies demonstrate that DF3 also binds directly to β-catenin and that the SXXXXXSSL motif in DF3 is responsible for

this interaction. Similar results were obtained with the highly related γ-catenin. Whereas the cytoplasmic domain of DF3/MUC1 is phosphorylated on tyrosine (8, 9), it is not known if tyrosine sites influence binding of catenins to the serine-rich motif. The formation of a complex between DF3 and β-catenin (or γ-catenin) may differ from those found in other β-catenin complexes. The interaction of E-cadherin or APC complexes to the cytoskeleton is mediated by binding of β-catenin to α-catenin (16). By contrast, there was little if any α-catenin in the complex of DF3 and β-catenin. Moreover, while E-cadherin forms a stable complex with β-catenin in suspension and adherent cells, the interaction of DF3 with β-catenin is detectable following cell adhesion. Similar findings were obtained for the interaction of DF3 and γ-catenin. These findings support a role for DF3 in the adhesion of cells and provide support for a novel interaction of DF3 with catenins.

REFERENCES

1. Kufe, D., Inghirami, G., Abe, M., Hayes, D., Justi-Wheeler, H., and Schlom, J. (1984) *Hybridoma* 3, 223–232
2. Siddiqui, J., Abe, M., Hayes, D., Shani, E., Yunis, E., and Kufe, D. (1988) *Proc. Natl. Acad. Sci. U. S. A.* 85, 2320–2323
3. Gendler, S. J., Burchell, J. M., Duhig, T., Lamport, D., White, R., Parker, M., and Taylor-Papadimitriou, J. (1987) *Proc. Natl. Acad. Sci. U. S. A.* 84, 6060–6064
4. Gendler, S., Taylor-Papadimitriou, J., Duhig, T., Rothbard, J., and Burchell, J. A. (1988) *J. Biol. Chem.* 263, 12820–12823
5. Abe, M., and Kufe, D. (1989) *Biochem. Biophys. Res. Commun.* 165, 644–649
6. Ligtenberg, M. J. L., Vos, H. L., Gennissen, A. M. C., and Hilkens, J. (1990) *J. Biol. Chem.* 265, 5573–5578
7. Wreschner, D. H., Hareuveni, M., Tsarfaty, I., Smorodinsky, N., Horev, J., Zaretsky, J., Kotkes, P., Weiss, M., Lathe, R., Dion, A., and Keydar, I. (1990) *Eur. J. Biochem.* 189, 463–473
8. Pandey, P., Kharbanda, S., and Kufe, D. (1995) *Cancer Res.* 55, 4000–4003
9. Zrihan-Licht, S., Baruch, A., Keydar, I., and Wreschner, D. H. (1994) *FEBS Lett.* 356, 130–137
10. Ligtenberg, M. J. L., Buijs, F., Vos, H. L., and Hilkens, J. (1992) *Cancer Res.* 52, 2318–2324
11. Wesseling, J., van der Valk, S. W., and Hilkens, J. (1996) *Mol. Biol. Cell* 7, 565–577
12. van de Wiel-van Kemenade, E., Ligtenberg, M. J. L., de Boer, A. J., Buijs, F., Vos, H. L., Melief, C. J. M., Hilkens, J., and Figdor, C. G. (1993) *J. Immunol.* 151, 767–776
13. Ozawa, M., Baribault, H., and Kemler, R. (1989) *EMBO J.* 8, 1711–1717
14. Ozawa, M., Ringwald, M., and Kemler, R. (1990) *Proc. Natl. Acad. Sci. U. S. A.* 87, 4246–4250
15. Rimm, D. L., Koslov, E. R., Kebriaei, P., Cianci, C. D., and Morrow, J. S. (1995) *Proc. Natl. Acad. Sci. U. S. A.* 92, 8813–8817
16. Hulsken, J., Birchmeier, W., and Behrens, J. (1994) *J. Cell Biol.* 127, 2061–2069
17. Knudsen, K. A., and Wheelock, M. (1992) *J. Cell Biol.* 118, 671–679
18. Peifer, M., McCrea, P. D., Green, K. J., Wieschaus, E., and Gumbiner, B. M. (1992) *J. Cell Biol.* 118, 681–689
19. McCrea, P. D., Turck, C. W., and Gumbiner, B. (1991) *Science* 254, 1359–1361
20. Noordermeer, J., Klingensmith, J., Perrimon, N., and Nusse, R. (1994) *Nature* 367, 80–83
21. Oda, H., Uemura, T., Shiomi, K., Nagafuchi, A., Tsukita, S., and Takeichi, M. (1993) *J. Cell Biol.* 121, 1133–1140
22. Oda, H., Uemura, T., Harada, Y., Iwai, Y., and Takeichi, M. (1994) *Dev. Biol.* 165, 716–726
23. Su, L. K., Vogelstein, B., and Kinzler, K. W. (1993) *Science* 262, 1734–1737
24. Rubinfeld, B., Souza, B., Albert, I., Muller, O., Chamberlain, S. H., Mastarz, F. R., Munemitsu, S., and Polakis, P. (1993) *Science* 262, 1731–1734
25. Rubinfeld, B., Souza, B., Albert, I., Munemitsu, S., and Polakis, P. (1995) *J. Biol. Chem.* 270, 5549–5555
26. Kintner, C. (1992) *Cell* 69, 225–236
27. Stappert, J., and Kemler, R. (1994) *Cell Adhes. Commun.* 2, 319–327
28. Hoschuetzky, H., Aberle, H., and Kemler, R. (1994) *J. Cell Biol.* 127, 1375–1380
29. Peifer, M., Berg, S., and Reynolds, A. B. (1994) *Cell* 76, 789–791

US005851775A

United States Patent [19]

Barker et al.

[11] **Patent Number: 5,851,775**

[45] **Date of Patent: Dec. 22, 1998**

[54] **β-CATENIN, TCF-4, AND APC INTERACT TO PREVENT CANCER**

[75] Inventors: **Nick Barker**, Utrecht; **Hans Clevers**, Ruysdaellaan, both of Netherlands; **Kenneth W. Kinzler**, Belair, Md.; **Vladimer Korinek**, Prague, Czech Rep.; **Patrice J. Morin**, Columbia, Md.; **Andrew B. Sparks**; **Bert Vogelstein**, both of Baltimore, Md.

[73] Assignees: **Johns Hopkins University**, Baltimore, Md.; **Utrecht University**, Netherlands

[21] Appl. No.: **821,355**

[22] Filed: **Mar. 20, 1997**

[51] **Int. Cl.**[6] **C12Q 1/68**; G01N 33/53

[52] **U.S. Cl.** **435/6**; 435/7.1; 435/189; 435/366

[58] **Field of Search** 435/6, 36, 7.1, 435/189, 366

[56] **References Cited**

PUBLICATIONS

Curt Suplee, The Washington Post, Key Process In Cancer Is Identified, Finding May Facilitate Early Detection, Therapy, 2–4 (1997).

Bonnee Rubinfeld et al. The Journal of Biological Chemistry, "The APC Protein and E–cadherin Form Similar but Independent Complexes with α–Catenin, β–Catenin, and Plakoglobin", vol. 270, No. 10, pp. 5549–5555 (1995).

J. Kawanishi et al., Molecular and Cellular Biology, "Loss of E–Cadherin–Dependent Cell–Cell Adhesion due to Mutation of the β–Catenin Gene in a Human Cancer Cell Line, HSC–39" vol. 15, No. 3, pp. 1175–1181 (1995).

P.F. Robins, J. Exp. Med., "A Mutated β–Catenin Gene Encodes a Melanoma–specific Antigen Recognized by Tumor Infiltrating Lymphocytes", vol. 183, pp. 1185–1192 (1996).

B. Rubinfeld et al., Science, "Association of the APC Gene Product with β–Catenin", vol. 262, pp. 1731–1734 (1993).

L.K. Su et al., Science, "Association of the APC Tumor Suppressor Protein with Catenins", vol. 262, pp. 1734–1737 (1993).

Bonnee Rubinfeld et al., Science, "Binding of GSK3β to the APC–β–Catenin Complex and Regulation of Complex Assembly", vol. 272, pp 1023–1026 (1996).

Susan Munemitsu et al., Proc. Natl. Acad. Sci., "Regulation of intracellular β–catenin levels by the adenomatous *polyposis coli* (APC) tumor–suppressor protein", vol. 92, pp. 3046–3050 (1995).

Jurgen Behrens et al., Nature, "Functional interaction of βcatenin with the transcription factor LEF–1" vol. 382, pp. 638–642 (1996).

Otmar Huber et al., Mech. Dev., "Nuclear localization of β–catenin by interaction with transcription factor LEF–1", vol. 59, pp. 3–10 (1996).

Miranda Molenaar, Cell, "XTcf–3 Transcription Factor Mediates β–Catenin–Induced Axis Formation in Xenopus Embryos", vol 86, (1996).

Primary Examiner—John L. LeGuyader
Assistant Examiner—Robert Schwartzman
Attorney, Agent, or Firm—Banner & Witcoff, Ltd.

[57] **ABSTRACT**

The APC tumor suppressor protein binds to β-catenin, a protein recently shown to interact with Tcf/Lef transcription factors. Here, the gene encoding a Tcf family member that is expressed in colonic epithelium (hTcf-4) was cloned and characterized. hTcf-4 transactivates transcription only when associated with β-catenin. Nuclei of $APC^{-/-}$ colon carcinoma cells were found to contain a stable β-catenin-hTCF-4 complex that was constitutively active, as measured by transcription of a Tcf reporter gene. Reintroduction of APC removed β-catenin from hTcf4 and abrogated the transcriptional transactivation. Constitutive transcription of TCF target genes, caused by loss of APC function, may be a crucial event in the early transformation of colonic epithelium. It is also shown here that the products of mutant APC genes found in colorectal tumors are defective in regulating β-catenin/Tcf-4 transcrpitional activation. Furthermore, colorectal tumors with intact APC genes were shown to contain subtle activating mutations of β-catenin that altered functionally significant phosphorylation sites. These results indicate that regulation of β-catenin is critical to APC's tumor suppressive effect and that this regulation can be circumvented by mutations in either APC or β-catenin.

9 Claims, 13 Drawing Sheets

EXHIBIT 3

FIG. 1A

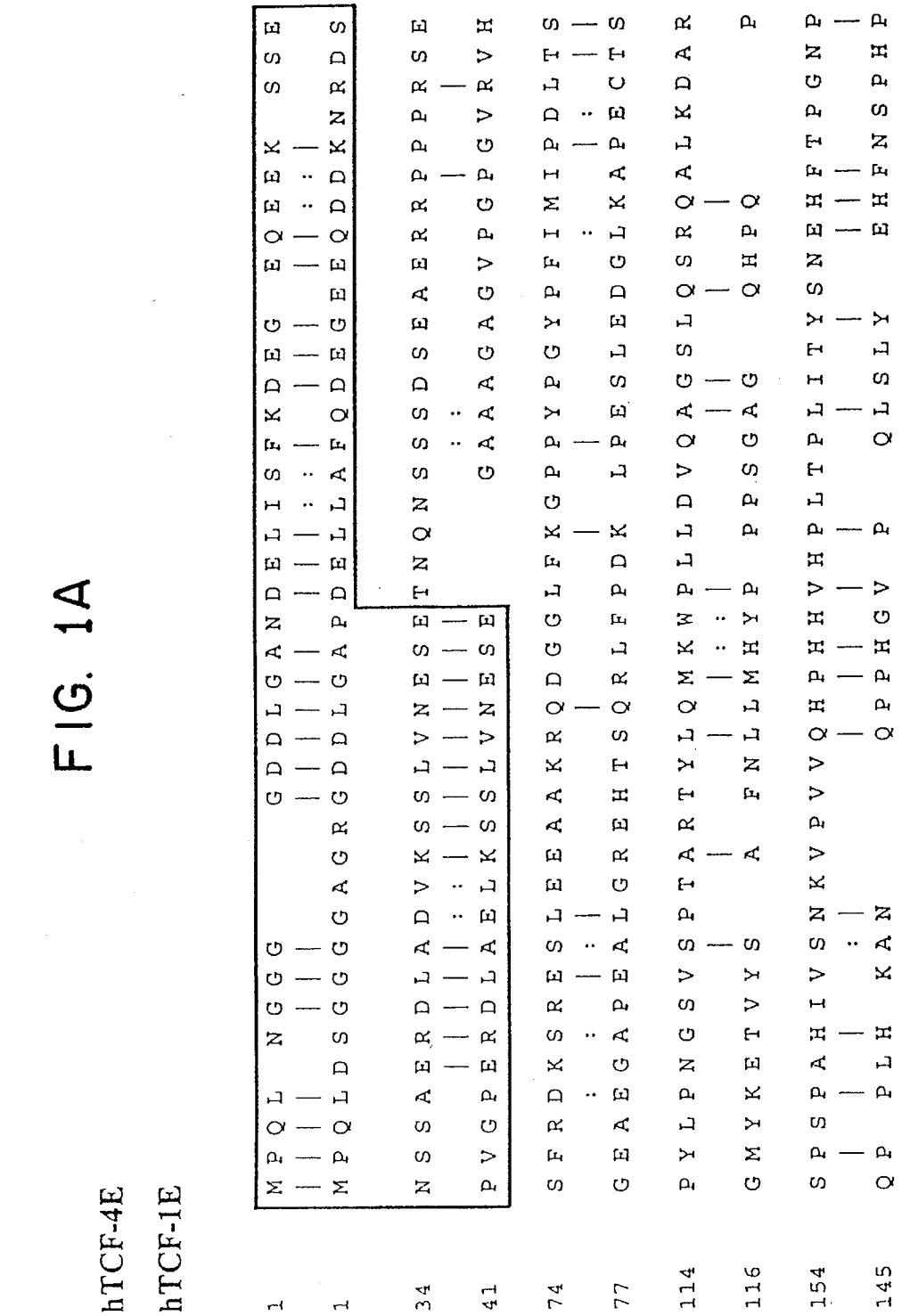

hTCF-4E
hTCF-1E
1 MPQLNGGG GDDLGANDELISFKDEG EQEEKSSE
1 MPQLDSGGGAGRGDDLGAPDELLAFQDEGEEQDDKNRDS
34 NSSAERDDLADVKSSLVNESETNQNSSSDSEAERRPPPRSE
41 PVGPERDDLAELKSSLVNESEGAAAGAGVPGGPGVRVH
74 SFRDKSRESLEEAAKRQDGGLFKGPPYPGYPPFIMIPDLTS
77 GEAEGAPEALGREHTSQRLFPDKLPESLEDGLKAPECTS
114 PYLPNGSVSPTARTYLQMKWPLLDVQAGSLQSRQALKDAR
116 GMYKETVYSAFNLLMHYPPPSGAGQHPQP
154 SPSPAHIVSNKVPVVQHPHHVHPLTPLITYSNEHFTPGNP
145 QPPLHKAN QPPHGVPQLSLYEHFNSPHP

FIG. 1B

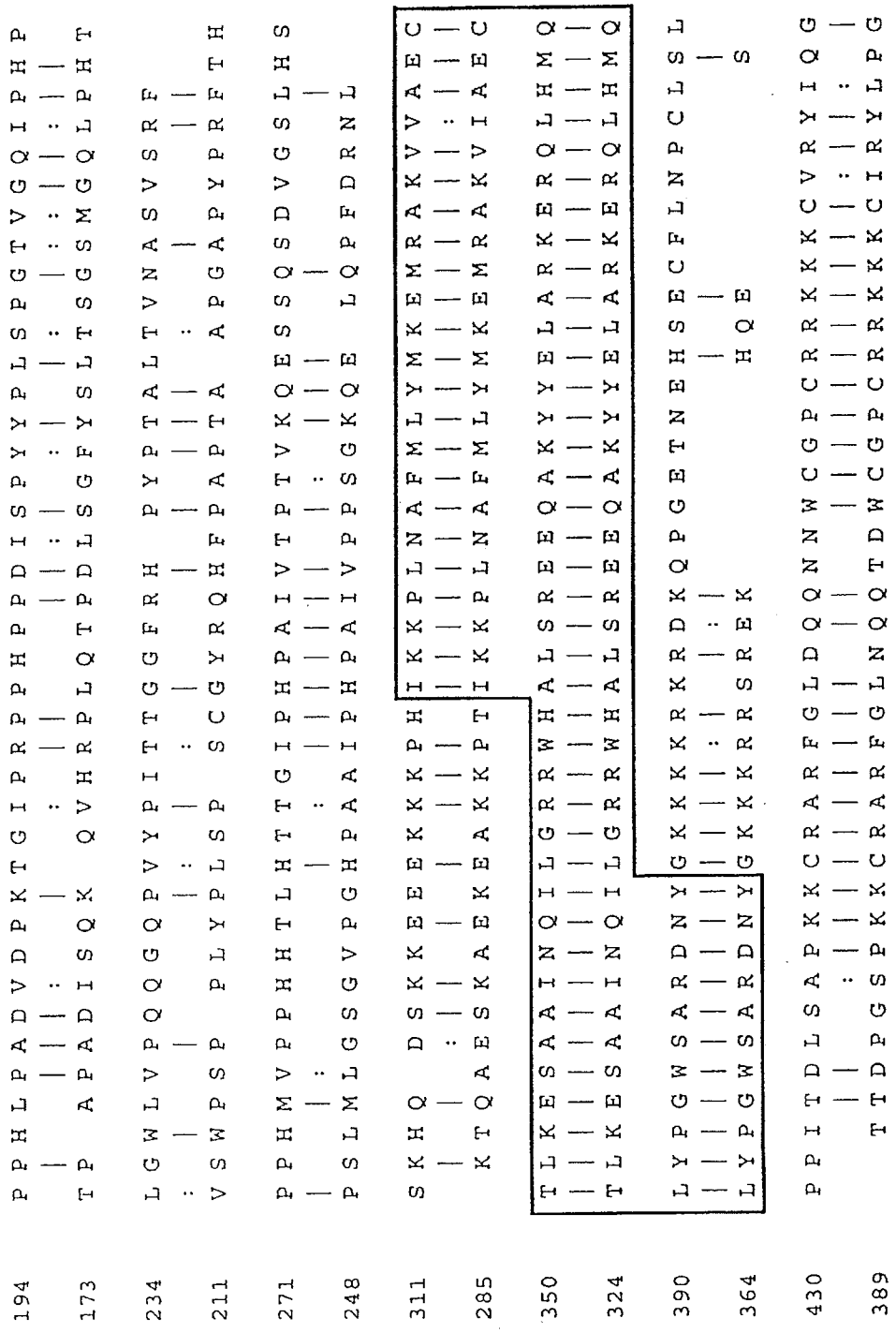

194 PPHLPADVDPKTGIPRPPPHPPDISPYYPLSPGTVGQIPHP
173 TP APADISQK QVHRPLQTPDLSGFYSLTSGSMGQLPHT
234 LGWLVPQQGQPVYPITTGGFRH PYPTALTVNASVSRF
211 VSWPSP PLYPLSP SCGYRQHFPAPTA APGAPYPRFTH
271 PPHMVPPHHTLHTTGIPHPAIVTPTVKQESSQSDVGSLHS
248 PSLMLGGSGVPGHPAAIPHPAIVPPSGKQE LQPFDRNL
311 SKHQ DSKKEEEKKKPHIKKPLNAFMLYMKEMRAKVVAEC
285 KTQAESKAEEAKKPTIKKPLNAFMLYMKEMRAKVIAEC
350 TLKESAAINQILGRRWHALSREEEQAKYYELARKERQLHMQ
324 TLKESAAINQILGRRWHALSREEEQAKYYELARKERQLHMQ
390 LYPGWSARDNYGKKKKRRDKQPGETNEHSECFLNPCLSL
364 LYPGWSARDNYGKKKRRSREK HQE S
430 PPITDLSAPKKCRARFGLDQQNNWCGPCRRKKKCVRYIQG
389 TTDPGSPKKCRARFGLNQQTDWCGPCRRKKKCIRYLPG

FIG. 1C

```
470   E G S C L S P P S S D G S L L D S P P P S P N L L G S P P R D A K S Q T E Q T Q
      | |   |   | |     | |   |   |     |   | |       |   |                   |
427   E G R C P S P V P S D D S A L   G C P G S P A P Q D S   P S Y H L L P R F P T E

510   P L S L S L K P D P   L A H L S     M M P P P P A L L L A E A T H K A S A L C P
        | :       |   |       | |     :     |   |       |     :       | :
465   L L T S P A E P A P T S P G L S T A L S L P T P G P P Q A P R S T L Q S T Q V Q

547   N G A L D L P P A A L Q P A A P S S S I A Q P S T S W L H S H S S L A G T Q P Q
      :               |
505   Q Q E S Q R Q V A *

587   P L S L V T K S L E *
```

hTCF-4B

hTCF-1B

```
390   L Y P G W S A R D N Y G K K K K R K R D K Q P G E T N G E K K S A F A T Y K V K
      | | | | | | | | | | | | | : |   | : |       |   |   | :   | |   | |     |
364   L Y P G W S A R D N Y G K K K R R S R E K H Q E S T T G G K R N A F G T Y P E K

430   A A A S A H P L Q M E A Y *
      | | |   |     |   |
404   A A A P A P F L P M T V L *
```

FIG. 2A

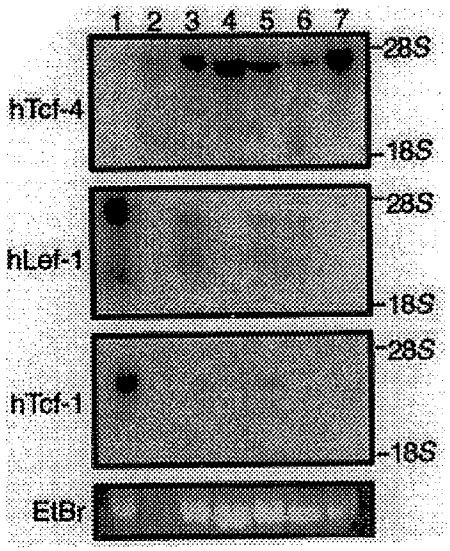


FIG. 2B

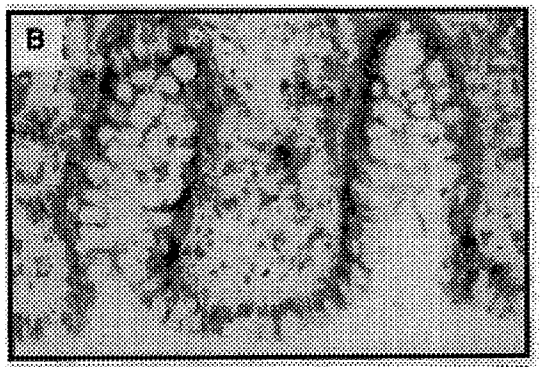


FIG. 2C

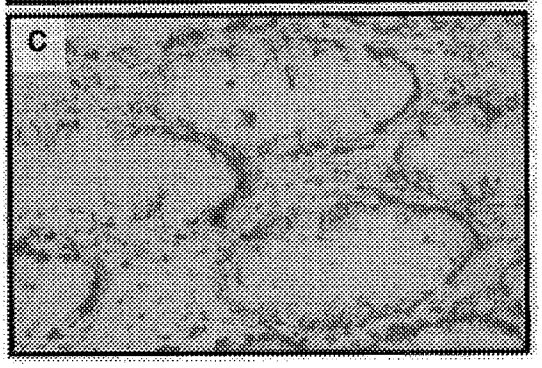


FIG. 3A

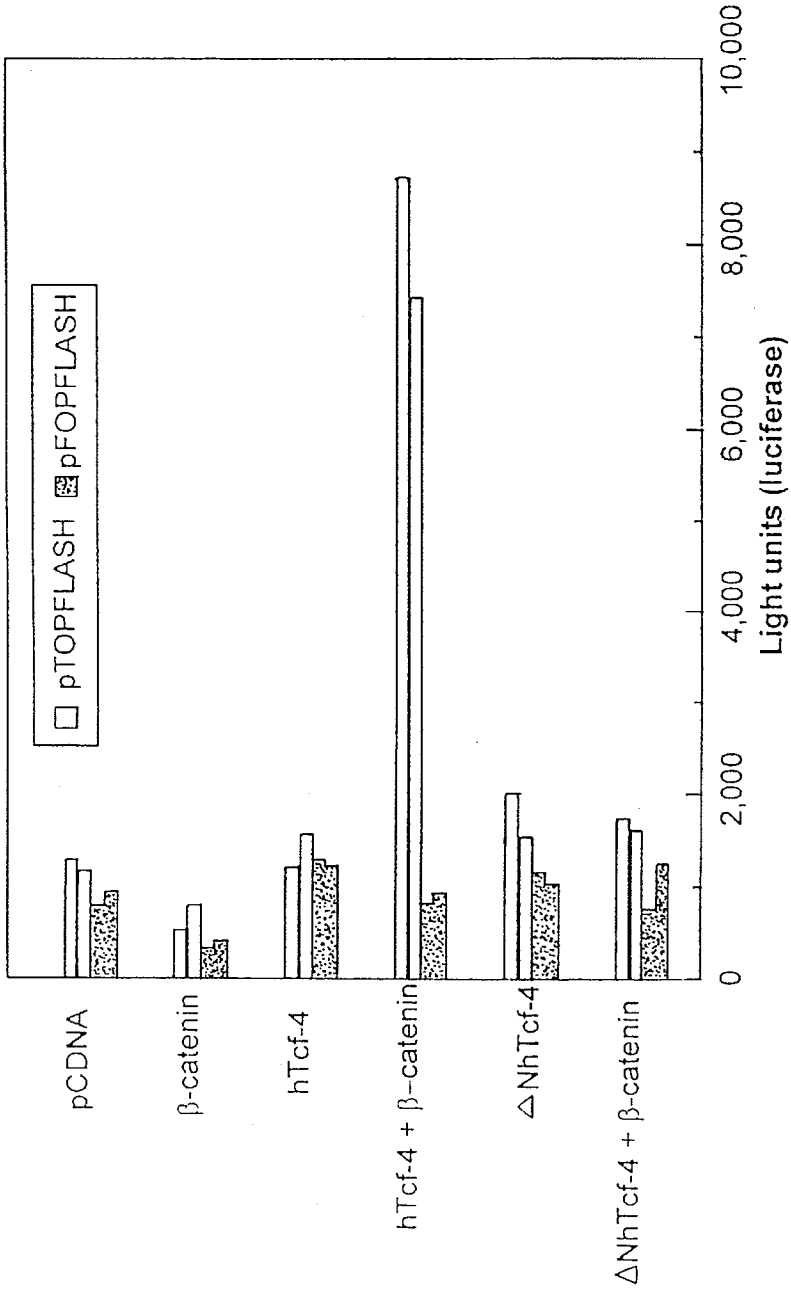

pTOPFLASH
pFOPFLASH
0
2,000
4,000
6,000
8,000
10,000
Light units (luciferase)
pCDNA
β-catenin
hTcf-4
hTcf-4 + β-catenin
ΔNhTcf-4
ΔNhTcf-4 + β-catenin

FIG. 3C

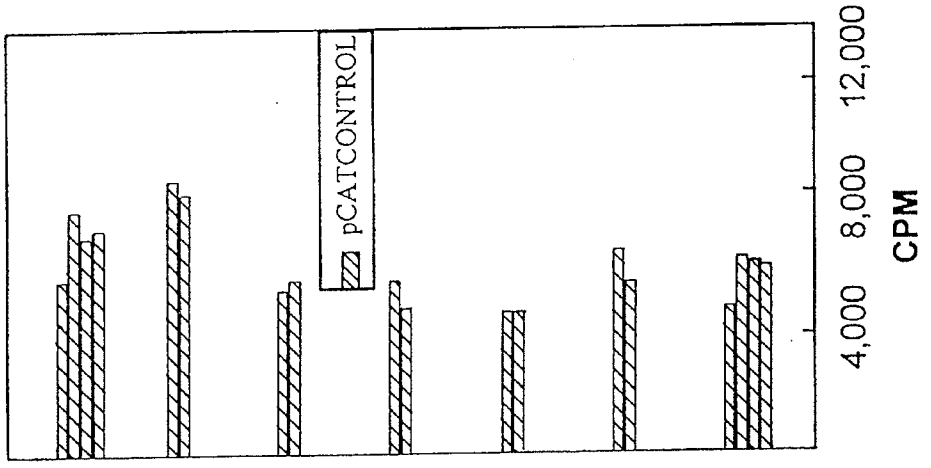

pCATCONTROL
4,000
8,000
12,000
CPM

FIG. 3B

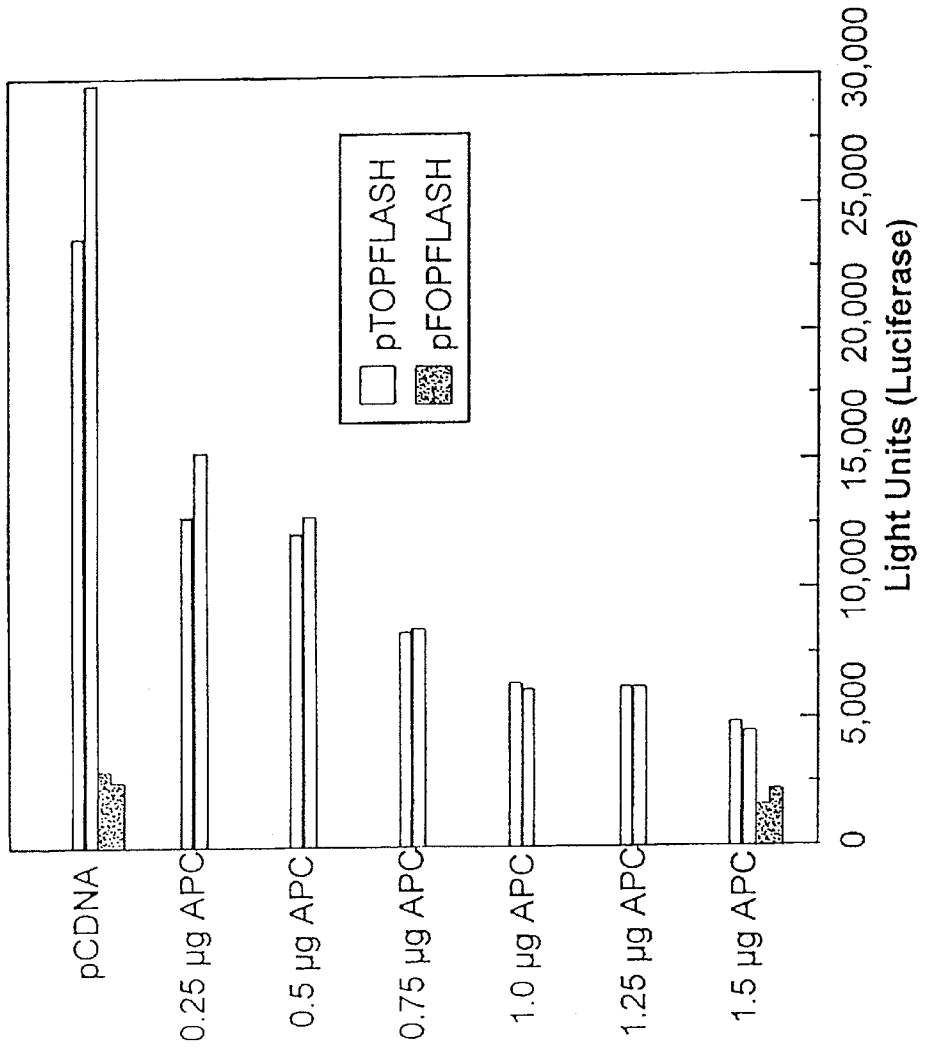

pTOPFLASH
pFOPFLASH
0
5,000
10,000
15,000
20,000
25,000
30,000
Light Units (Luciferase)
pCDNA
0.25 µg APC
0.5 µg APC
0.75 µg APC
1.0 µg APC
1.25 µg APC
1.5 µg APC

FIG. 4

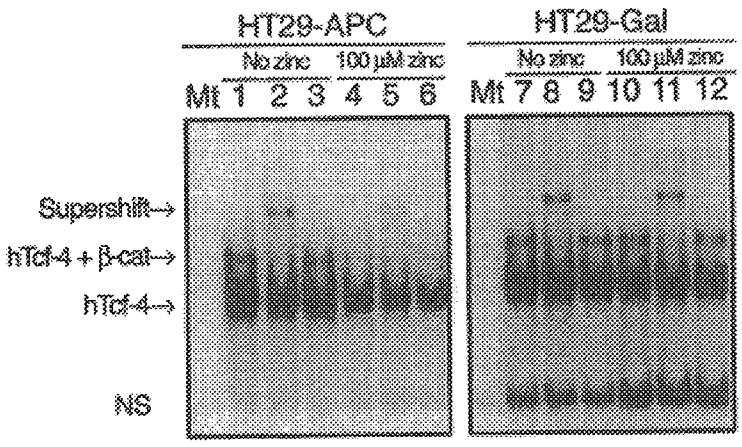

HT29-APC
HT29-Gal
No zinc
100 μM zinc
No zinc
100 μM zinc
Mt 1 2 3 4 5 6
Mt 7 8 9 10 11 12
Supershift→
hTcf-4 + β-cat→
hTcf-4→
NS

FIG. 5A

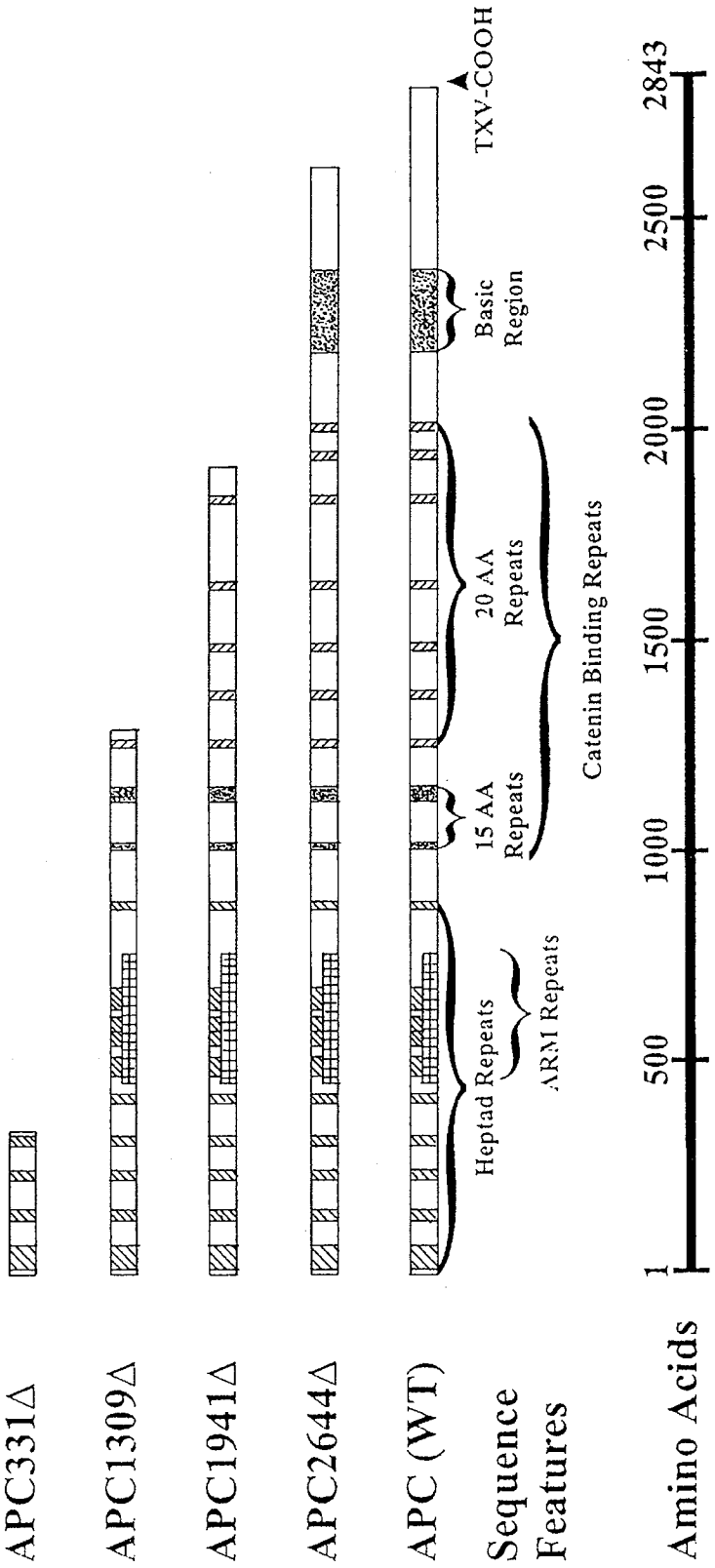

APC331Δ
APC1309Δ
APC1941Δ
APC2644Δ
APC (WT)
Sequence Features
Amino Acids
Heptad Repeats
ARM Repeats
15 AA Repeats
20 AA Repeats
Catenin Binding Repeats
Basic Region
TXV-COOH
1
500
1000
1500
2000
2500
2843

FIG. 5B

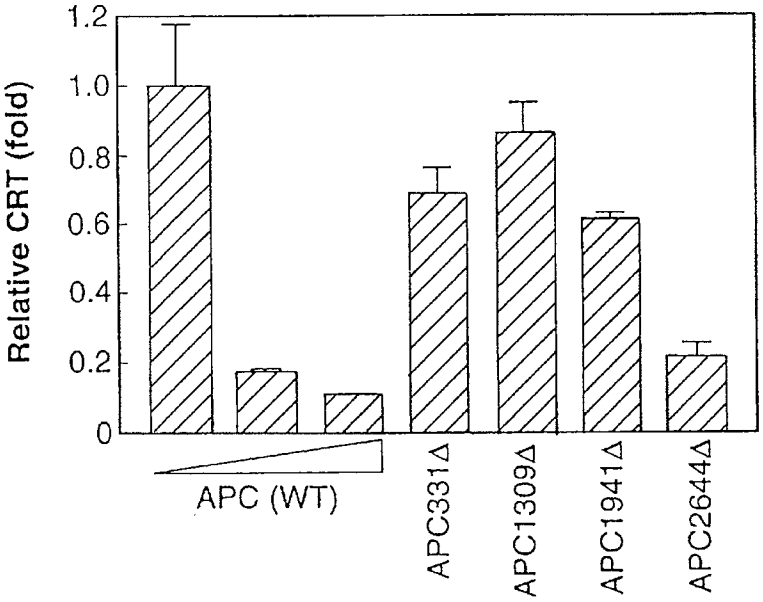

Relative CRT (fold)
1.2
1.0
0.8
0.6
0.4
0.2
0
APC (WT)
APC331Δ
APC1309Δ
APC1941Δ
APC2644Δ

FIG. 6B

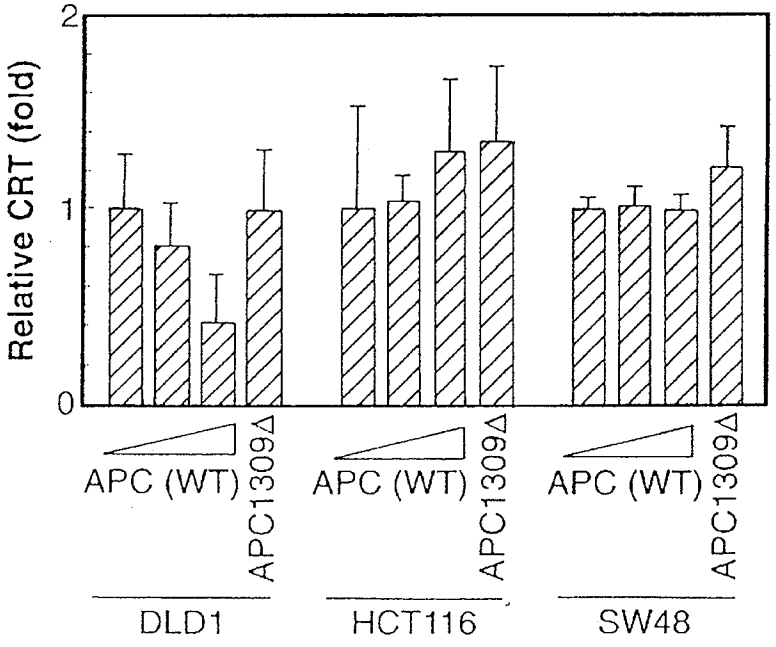

Relative CRT (fold)
2
1
0
APC (WT)
APC1309Δ
APC (WT)
APC1309Δ
APC (WT)
APC1309Δ
DLD1
HCT116
SW48

FIG. 6A

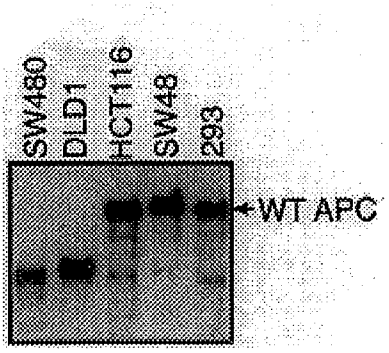

SW480
DLD1
HCT116
SW48
293
WT APC

FIG. 7A

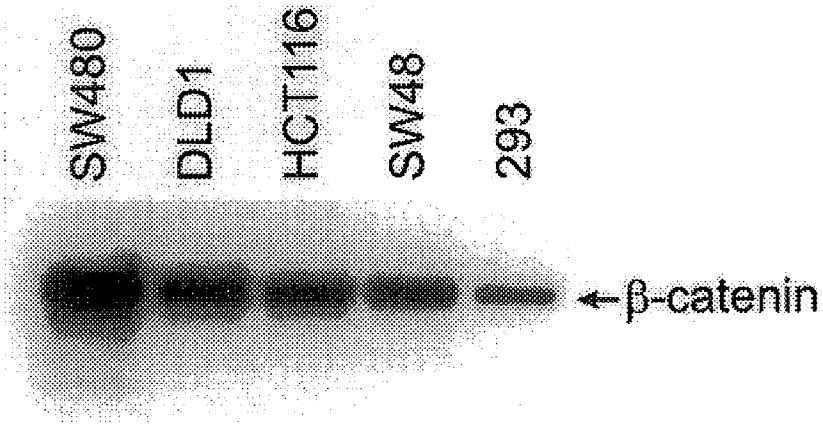

SW480
DLD1
HCT116
SW48
293
β-catenin

FIG. 8A

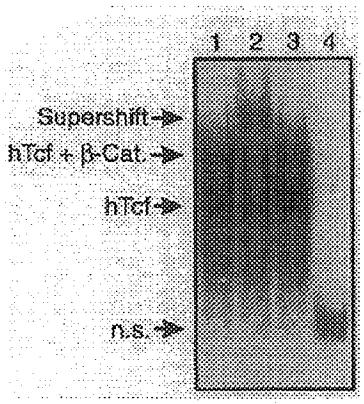

1 2 3 4
Supershift
hTcf + β-Cat.
hTcf
n.s.

FIG. 7B

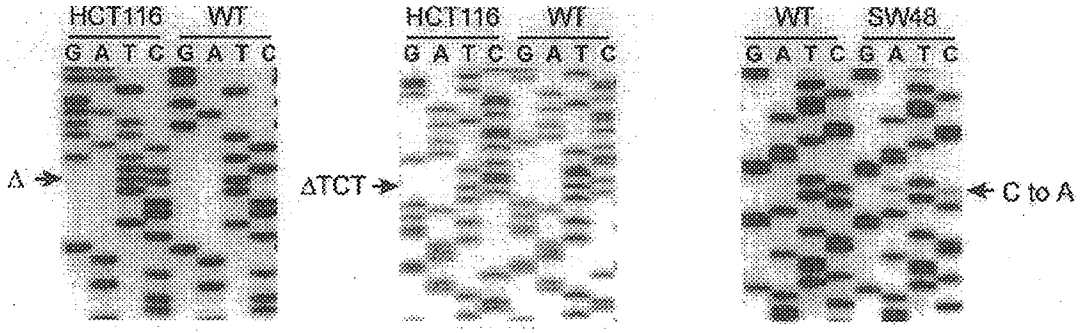

HCT116 WT
G A T C G A T C
A
HCT116 WT
G A T C G A T C
ΔTCT
WT SW48
G A T C G A T C
C to A

FIG. 7C

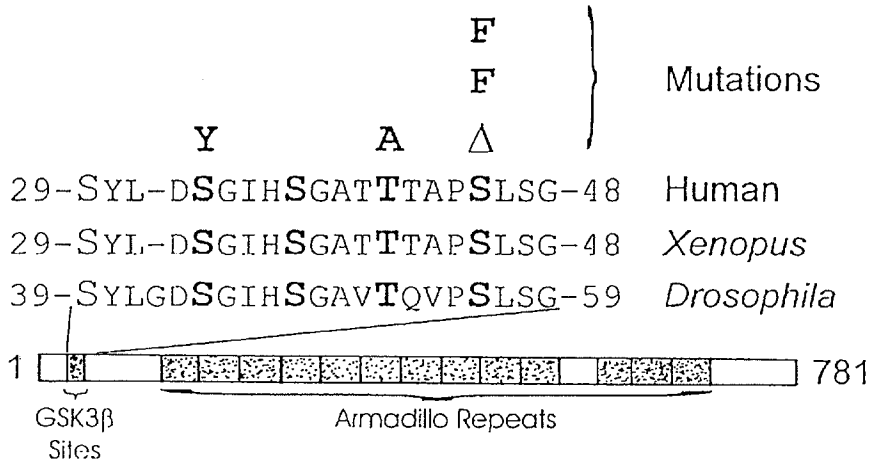
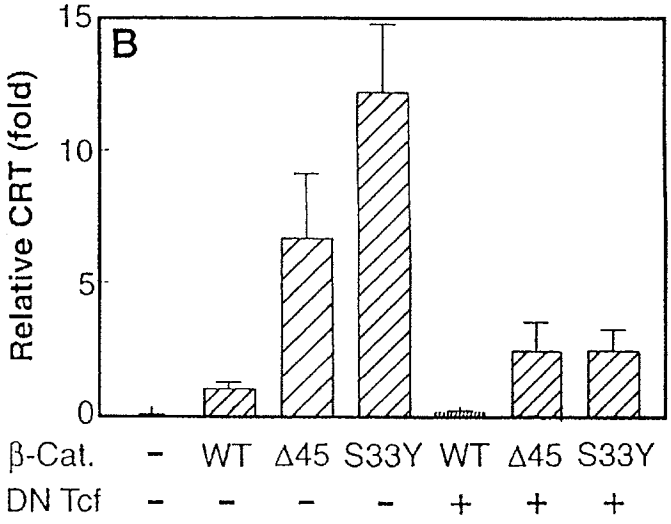

F
F
Y A Δ
Mutations
29-SYL-DSGIHSGATTTAPSLSG-48 Human
29-SYL-DSGIHSGATTTAPSLSG-48 Xenopus
39-SYLGDSGIHSGAVTQVPSLSG-59 Drosophila
1
781
GSK3β Sites
Armadillo Repeats

FIG. 8B



β-CATENIN, TCF-4, AND APC INTERACT TO PREVENT CANCER

The U.S. Government has a paid-up license in this invention and the right in limited circumstances to require the patent owner to license others on reasonable terms as provided for by the terms of grant CA57345 awarded by the National Institutes of Health.

TECHNICAL FIELD OF THE INVENTION

This invention is related to the field of cancer diagnostics and therapeutics. More particularly it relates to methods for diagnosing and treating cancers associated with APC or β-catenin mutations.

BACKGROUND OF THE INVENTION

Mutations of the adenomatous polyposis coli (APC) gene are the most common disease-causing genetic events in humans; approximately 50% of the population will develop colorectal polyps initiated by such mutations during a normal life span (14). Individuals who inherit APC mutations develop thousands of colorectal tumors, consistent with APC's tumor suppressor or "gatekeeping" role in colorectal tumorigenesis (15,16). APC homodimerizes through its amino-terminus (17), and interacts with at least six other proteins: β-catenin (18), γ-catenin (plakoglobin) (19), tubulin (20), EB1 (21), hDLG, a homologue of a Drosophila tumor suppressor protein (22), and ZW3/GSK3β kinase (23). Whether any of these interacting proteins communicate APC growth-controlling signals is unknown. Thus there is a need in the art for a fuller understanding of how the tumor suppressor gene APC functions in cells.

SUMMARY OF THE INVENTION

It is an object of the present invention to provide human nucleotide sequences encoding transcriptional activation proteins.

It is another object of the present invention to provide isolated preparations of transcriptional activation proteins.

It is an object of the present invention to provide methods of determining the presence or absence in a cell of wild-type APC or a downstream protein in the APC transcription regulatory pathway.

Another object of the invention is to provide methods of identifying candidate drugs for use in Familial Adenomatous Polyposis (FAP) patients or patients with increased risk of developing cancer.

It is yet another object of the invention to provide methods of identifying candidate drugs for the treatment of cancer patients, in particular those with APC or β-catenin mutations.

Another object of the invention is to provide a method for diagnosing cancer in a sample suspected of being neoplastic.

Another object of the invention is to provide a method for treating a patient with colorectal cancer or other cancer associated with FAP.

These and other objects of the invention are achieved by providing one or more of the embodiments described below. In one embodiment of the invention an intron-free DNA molecule is provided which encodes Tcf-4 protein as shown in SEQ ID NO: 5 or 6.

According to another embodiment of the invention an isolated Tcf-4 protein is provided. The protein is substantially free of other human proteins, and has a sequence as shown in SEQ ID NO: 5 or 6.

In another embodiment of the invention a method is provided for determining the presence or absence in a cell of wild-type APC or a downstream protein in the APC transcription regulatory pathway. The method comprises the steps of:

introducing a Tcf-responsive reporter gene into the cell; and

measuring transcription of said reporter gene; wherein a cell which supports active transcription of said reporter gene does not have wild-type APC or does not have a wild-type downstream protein in the APC transcription regulatory pathway.

According to yet another embodiment of the invention a method is provided for determining the presence or absence in a cell of wild-type APC. The method comprises the steps of:

contacting a Tcf-responsive reporter gene with a lysate of the cell; and

measuring transcription of said reporter gene; wherein a lysate which inhibits said transcription has wild-type APC.

In still another embodiment of the invention a method of identifying candidate drugs is provided. The drugs may be useful for treatment of FAP or other cancer patients or patients with increased risk of developing cancer. The method comprises the steps of:

contacting a cell having no wild-type APC or a mutant p-catenin with a test compound;

measuring transcription of a Tcf-responsive reporter gene, wherein a test compound which inhibits the transcription of the reporter gene is a candidate drug for cancer therapy.

According to yet another aspect of the invention another method is provided for identifying candidate drugs for use in for use in FAP patients, colon cancer patients, patients with mutations in β-catenin or APC, or patients with increased risk of developing cancer. The method, comprises the steps of:

contacting a Tcf-responsive reporter gene with a test compound under conditions in which the reporter gene is transcribed in the absence of the test compound; and

measuring transcription of the Tcf-responsive reporter gene; wherein a test compound which inhibits said transcription is a candidate drug for cancer therapy.

According to another aspect of the invention a method is provided for identifying candidate drugs for use in FAP patients or patients with increased risk of developing cancer. The method comprises the steps of:

contacting a test compound with β-catenin and Tcf-4 under conditions in which β-catenin and Tcf-4 bind to each other; and

determining whether the test compound inhibits the binding of β-catenin and Tcf-4, a test compound which inhibits the binding being a candidate for cancer therapy or prophylaxis.

According to still another embodiment of the invention a method is provided for diagnosing cancer in a sample suspected of being neoplastic. the method comprises the steps of:

comparing a CTNNB sequence found in the sample to a second CTNNB sequence found in a normal tissue, wherein a difference between the first and second sequence is an indicator of cancer.

According to another aspect of the invention a method is provided for treating a patient with colorectal cancer or other cancer associated with FAP.

The method comprises the step of:

administering to the patient a nucleotide sequence comprising a portion of the APC coding sequence, said portion consisting of the p-catenin binding site.

According to another aspect of the invention a method is provided for treating a patient with colorectal cancer or other cancer associated with FAP. The method comprises the step of:

administering to the patient a polypeptide comprising a portion of the APC coding sequence, said portion consisting of the β-catenin binding site.

The present invention thus provides the art with diagnostic, therapeutic and drug discovery methods especially useful for FAP and other cancers with APC or β-catenin mutations.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. **1**. Sequence comparison of hTcf-4 and hTcf-1.

Two alternative splice forms of hTcf-4 were identified, each encoding a different COOH-terminus. One form (hTcf-4E; SEQ ID NO:6) was homologous to hTCF-IE; SEQ ID NO:9(FIG. 1A) (7); the other form (hTcf-4B; SEQ ID NO: 5) was homologous to hTcf-1B; SEQ ID NO:8. (FIG. **1B**). The highly conserved NH_2-terminal interaction domain and the High Mobility Group (HMG) box DNA-binding region are boxed. Abbreviations for the amino acids are: A. Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; IC, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; P, Ar g; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

FIG. **2**. Analysis of hTcf-4 expression in colonic epithelium.

(FIGS. 2A, 2B, and 2C) Northern blot analysis of hTcf-4, hTcf-1, hLef-I expression in Jurkat T cells (lane 1); colonic mucosa (lane 2); colon carcinoma cell lines DLD-1 (lane 3), HCT116 (lane 4); SW480 (lane 5); SW620 (lane 6); HT29 (lane 7). Lane 2 contains 5 μg total RNA; all others contain 15 μg total RNA. The positions of 18S and 28S ribosomal RNAs are shown. EtBr, ethidium bromide stain. (FIG. **2B**) In situ hybridization of healthy human colon tissue to an hTcf-4 probe. (FIG. **2C**) In situ hybridization to a negative control probe (a fragment of the *E. coli* neomycin resistance gene).

FIGS. **3**A, **3**B. Transactivational properties of β-catenin/hTcf-4.

All reporter assays were performed as duplicate transfections. For each condition, both values are shown. (FIG. **3**A) Reporter gene assays in IIA1.6 B cells. Cells were transfected by electroporation with 1 μg luciferase reporter plasmid, 5 μg β-catenin expression plasmid, and 3 II-hTcf-4 expression plasmids. Empty pCDNA was added to a total of 10 μg, plasmid DNA. (FIG. **3**B) Reporter gene assays in SW480 colon carcinoma cells. Cells were transfected with 0.3 μg, of the indicated luciferase reporter gene, 0.7 μg pCATCONTROL as internal control, the indicated amounts of pCMVNeoAPC, and empty PCDNA to a total of 2.5 μg plasmid DNA. Control CAT values are given in the right panel.

FIG. **4**. Constitutive presence of β-catenin-hTcf-4 complexes in $APC^{-/-}$ cells. Gel retardation assays were performed on nuclear extracts from the indicated cell lines before and after a 20-hour exposure to Zn^{++}. Samples in lanes 1, 4, 7, 10 were incubated under standard conditions. To the samples in lanes 2, 5, 8, 11, 0.25 μg, anti β-catenin was added. To the samples in lanes 3, 6, 9, 12, 0.25 μg of a control (human CD4) antibody was added. N.S., nonspecific band also observed with mutant (nonbinding) probe (lane Mt).

FIGS. **5**A and **5**B. Effects of APC mutations on CRT. (FIG. 5A) Schematics of wild-type (WT) and mutant APC. APC is a 2843-amino-acid (AA) protein (32) with contains armadillo (ARM) repeats in the amino-terminus (33), 15 and 20 AA β-catenin-binding repeats in the central region (18, 19, and a basic region in the carboxyl-terminus (32). The carboxyl-terminus also contains a TXV sequence which mediates DLG binding (22). (FIG. 5B) Effects of WT and mutant APC on CRT. SW480 cells containing endogenous mutant APC were transfected with the APC expression vectors shown in (FIG. 5A) and CRT was measured. Cells were transfected with increasing amounts of WT APC (0, 0.15 and 0.5 μg) or 0.5 μg mutant APC. CRT reporter activities are expressed relative to assays containing no WT APC and are the means of three replicates. Error bars represent standard deviations.

Lipofectamine was used to cotransfect SW480 cells with an internal control (0.5 μg pCMV-βgal), a reporter construct (0.5 μg pTOPFLASH or pFOPFLASH) and the indicated amount of the various APC expression vectors. The pTOPFLASH reporter contained an optimized Tcf-binding site 5' of a luciferase reporter gene, whereas pFOPFLASH contained a mutated site that does not bind Tcf. The amount of DNA in each transfection was kept constant by addition of an appropriate amount of empty expression vector (pCEP4). Luciferase and β-galactosidase activities were determined 16 hours after transfection. Luciferase activity was corrected for transfection efficiency (using the control β-galactosidase activity) and nonspecific transcription (using the pFOPFLASH control).

FIGS. **6**A and **6**B. Evaluation of CRT in colorectal cancer cell lines with WT APC. (FIG. 6A) Immunoblot of endogenous APC in the DLD1, SW480, HCT116, SW48 and 293 cell lines, developed with APC monoclonal antibody FE9 (34). (FIG. 6B) Effects of exogenous WT APC on CRT in cell lines with endogenous mutated or WT APC. Cells were transfected with increasing amounts (0, 0.15 μg, 0.5 μg for DLD1 and SW48; 0, 0.5 μg, 5 ,g for HCT116) of WT APC or APC1309Δ mutant (0.5 μg for DLD1 and SW48; 5 μg for HCT116) and CRT was assessed as in FIG. **5**. CRT reporter activities are expressed relative to activity in extracts without exogenous APC and are the means of three replicates. Error bars represent standard deviations.

FIGS. **7**A, **7**B and **7**C. Evaluation of β-catenin in colorectal cancer cell lines with WT APC. (FIG. 7A) Immunoblot of the cell lines used in this study, developed with β-catenin monoclonal C19220 (Transduction Laboratories, Lexington, Ky.)(31). (FIG. 7B) Sequence of CTNNB1 in HCT116 and SW48. Overlapping segments constituting the entire CTNNB1 were amplified by RT-PCR from SW480, DLD1, HCT116, and SW48 cells, and sequenced directly with ThermoSequenase (Amersham). In the case of HCT116, a PCR product containing the deleted region was also cloned into pCI-neo (Promega, Madison) and multiple clones corresponding to each allele were individually sequenced.

The left panel (nts 121 to 143 from HCT116) reveals the presence of a deletion in addition to the WT sequence. The middle panel (antisense strand 156 to 113 of the WT and deleted alleles of HCT116) reveals the 3-bp deletion (ΔTCT) that removed codon 45 in half the clones. The right panel (nts 80 to 113 from SW48) reveals a C to A transition affecting codon 33 (TCT to TAT). FIG. 7C) Schematic of β-catenin illustrating the armadillo repeats (33) in human (SEQ ID NO:10), Xenopus (SEQ ID NO:10) and drosophile (SEQ ID NO:11) and negative regulatory domain. The residues in larger type fit the consensus sequence for GSK3β

phosphorylation (29) and those in bold have been demonstrated to affect down regulation of β-catenin through GSK3β phosphorylation in Xenopus embryos (27). The five mutations found in human colon cancers are indicated at the top.

FIGS. 8A and B. Functional evaluation of β-catenin mutants. (FIG. 8A) Constitutive nuclear complex of β-catenin and Tcf in HCT116 cells. The presence of nuclear β-catenin-Tcf complexes was assessed by gel shift assays. Lanes 1 to 3, optimal Tcf retardation probe shifted with nuclear extract from HCT116 cells with addition of no antibody (lane 1), anti β-catenin (0.25 μg, lane 2), or an irrelevant antibody (0.25 μg, lane 3). Lane 4, mutant Tcf retardation probe shifted with nuclear extract from HCT1116 cells. n.s., nonspecific shifting seen with the mutant probe. (FIG. 8B) Effects of the β-catenin mutations on CRT. 293 cells were transfected with WT (WT) or mutant (Δ45, S33Y) β-catenin and CRT was assessed. CRT reporter activities are expressed relative to WT β-catenin and are the means of three replicates. Error bars represent standard deviations. β-catenin expression constructs were prepared as follows. WT CTNNB1 was amplified by RT-PCR from SW480 cells and cloned into the mammalian expression vector pCI-neo (Promega) to produce pCI-neo-β-cat. The pCI-neo-β-cat Δ45 and S33Y were generated by replacing codons 1 to 89 in pCI-neo-β-cat with a PCR product encoding the equivalent region from HCT116 or SW48 cDNA, respectively. The structures of all constructs were verified by sequence analysis. Lipofectamine was used to cotransfect 293 cells with an internal control (0.1 μg CMV-βgal), a reporter (0.5 μg pTOPFLASH or pFOPFLASH), a Tcf-4 expression vector (0.5 μg pCDNA-TCF4), and β-catenin (0.5 μg) or dominant negative hTcf-4 1.0 μg) expression vectors. CRT was determined as described above.

DETAILED DESCRIPTION

It is a discovery of the present invention that hTcf-4 binds to β-catenin and activates transcription in colorectal epithelial cells. Moreover, it has now been found that APC regulates this transcriptional activation, at least in part by binding to β-catenin. In colorectal cancer cells this regulation is frequently abrogated, either by mutation of APC or by mutation of β-catenin.

Two alternative splice forms of human Tcf-4 have been found. One form (hTcf-4E) is homologous to hTcf-1E and the other (hTcf-4B) is homologous to hTcf-1B. The sequence of the nucleotide and amino acid sequences are shown in SEQ ID NOs: 1, 2, 5, and 6. The coding sequences and proteins can be used in assays as described below. Intron-free DNA molecules are provided which are originally made by reverse transcription of a mRNA molecule. They can be propagated in cells or amplified as is desired. Isolated Tcf-4 proteins can be provided substantially free of other human proteins if, for example, the nucleotide sequences are expressed in non-human cells. Methods and vectors for achieving such expression are well known in the art. Choice of such expression means is made by the skilled artisan according to the desired usage and convenience.

Cells can be tested to determine if they have a wild-type APC or a wild-type downstream protein in the APC transcription regulatory pathway, called herein the CRT pathway (β-catenin/Tcf-regulated transcription). One protein within the CRT pathway which has been identified as a target of mutations in human cancers is β-catenin (encoded by the CTNNB1 gene). Other parts of the pathway are also likely to be targets. Although the target genes of the CRT pathway have not been identified, they can be readily identified using the system disclosed here. Genes which are differentially transcribed in the presence of wild-type and mutant CTNNB1, for example, can be identified.

Tcf-responsive reporter genes are those constructs which comprise a readily detectable or assayable gene (such as luciferase, β-galactosidase, chloramphenicol acetyltransferase) linked in cis to a Tcf-responsive element. Such responsive elements are known in the art (7) and any such elements can be used. An optimal Tcf motif contains the sequence CCTTTGATC. From one to twenty copies, and preferably from three to six copies, of the motif may be used. Mutation of the sequence to CCTTTGGCC abrogates responsiveness. Another necessary part of such constructs is a minimal promoter, such as the c-Fos or the Herpes virus thymidine kinase promoter. Transcription of the reporter gene may be performed by any means known in the art, usually by assaying for the activity of the encoded gene, although immunological detection methods can also be used. In addition, transcription can be monitored by measuring the transcribed mRNA directly, typically using oligonucleotide probes.

As shown below, a cell which has a wild-type APC protein will inhibit CRT. However, most mutations in APC render APC unable to inhibit CRT. Similarly, certain mutations in CTNNB1 render β-catenin super-active and/or refractory to the inhibition by APC. Thus measuring Tcf-responsive reporter gene transcription is an indication of the status of APC and CTNNB1. Mutations in both of these genes are associated with cancers and therefore provides diagnostic and prognostic information.

Assays for CRT can be accomplished in vitro or in cells. If the assay is to be accomplished in cells, then a Tcf-responsive reporter gene must be introduced into the cell. Any means for introducing genetic material into cells can be used, including but not limited to infection, transfection, electroporation. If the assay is to be performed in vitro then the components for transcription must be present. These include suitable buffers, RNA polymerase, as well as ribonucleotides. If the protein product is to be assayed, then the components for translation must also be present, such as ribosomes, and amino acids.

These assays can also be used to screen compounds for potential as anti-cancer therapeutic agents. Using either the in vitro or cell form of the assay, test compounds can be introduced to determine whether they are able to mimic the effect of wild-type APC or to convert a mutant APC into a form which is able to inhibit CRT or a mutant β-catenin into a form which is regulatable by APC. In addition, compounds can be tested for the ability to inhibit the binding of β-catenin and Tcf-4, thus mimicking the action of APC. Such a test can be conducted in vitro or in vivo, for example using a two hybrid assay.

A means for diagnosis of cancers is the result of the observation that CTNNB1 mutations are found in tumor cells, especially those which have wild-type APC. Such mutations can be found, inter alia, by sequencing either the gene or the protein found in a sample. Functional assays can also be used, such as whether β-catenin binds to APC or Tcf-4, or whether it is capable of mediating CRT. Sequences can be compared to those found in a normal tissue of a human, especially the same human who provided the sample being tested. Suitable tumors for testing include, but are not limited to those which are associated with FAP. Suitable tumors include colorectal cancer, thyroid cancer, brain

cancer, medulloblastoma, desmoid tumor, osteoma, breast cancer, and head and neck cancer. Because APC mutations are so frequent, and because it appears that APC mutations do not occur in the same tumors as CTNNB1 mutations, one can prescreen samples for APC mutations before performing a CTNNB1 determination.

The portion of the APC gene which encodes the β-catenin binding site can be used in a gene therapy format. Suitable techniques are known in the art for administering genes to tumors, and any such technique can be used. Suitable expression vectors are also known in the art and it is within the skill of the artisan to select an appropriate one. Upon expression in a tumor cell of the β-catenin binding portion of APC, β-catenin will be bound and titrated away from binding to Tcf-4, thus preventing unregulated expression of the CRT target genes. Similarly, a polypeptide portion of APC containing the β-catenin binding site can be administered to cells to perform a titration of β-catenin. Techniques for such administration to cells is well known in the art. Cells which are treated with either the polynucleotide or the polypeptide can be used to study the interaction between APC and β-catenin, and for developing drugs which interfere with such binding.

The above disclosure generally describes the present invention. A more complete understanding can be obtained by reference to the following specific examples which are provided herein for purposes of illustration only, and are not intended to limit the scope of the invention.

EXAMPLE 1

This example identifies Tcf-4 as the expressed family member in colorectal epithelial cells and provides the complete sequence of the cloned cDNA.

There are four known members of the Tcf/Lef family in mammals: the lymphoid-specific factors Tcf- I and Lef- 1 (7,8), and the less well characterized Tcf-3 and 4(9). We performed a qualitative Reverse Transcriptase-Polymerase Chain Reaction assay for expression of the four Tcf/Lef genes on 43 colon tumor cell lines. While most colon cell lines expressed more than one of the genes, only hTcf-4 mRNA was expressed in essentially all lines.

We then screened a human fetal CDNA library and retrieved clones encoding full-length hTcf-4 (FIG. 1). A genomic fragment encoding, the HMG box region of hTcf-4 (7) was used to probe a human 12 week-fetal cDNA library in Lambda GT-11. Positive clones were subcloned into pBluescriptSK and sequenced. See SEQ ID NOs: 1 and 2. The predicted sequence of hTcf-4 was most similar to that of hTcf-1. Alternative splicing yielded two COOH-termini that were conserved between hTcf- I and hTcf-4. The NH_2-terminus, which in hTcf- 1, mLef-1 and Xenopus TCF-3 mediates binding to β-catenin (6), was also conserved in hTcf-4. Northern blot analysis of selected colon carcinoma cell lines revealed high-level expression of hTcf-4 (FIG. 2A). Northern blot hybridizations (7) were performed with full-length hTcf-1, hLef-I and hTcf-4 cDNA. Colon epithelial cells were freshly prepared from a mucosal preparation dissected from a healthy surgical colon sample. The sample was minced, and incubated with 1 mM dithiothreitol (DTT) in Hanks' medium to remove mucus. Single-cell suspensions were prepared by incubation at RT in 0.75 mM EDTA in Hanks' medium. Epithelial cells were separated from lymphocytes by Percoll gradient centrifugation.

As evidenced by in situ hybridization (FIG. 2, B and C) and Northern blotting (FIG. 2A), hTcf-4 MRNA was readily detectable in normal colonic epithelium, whereas hTcf-I and hLef-I were not detectable. In situ hybridization of 6μ frozen sections of healthy colon biopsy samples were performed as described(10). hTcf-4 cDNA encoding amino acids 200 to 310 was amplified and labeled with Dig-11-dUTP (Boehringer Mannheim, Germany) by PCR. After hybridization and washing, the sections were sequentially incubated with mouse anti-Dig antibody (Boehringer) and a horseradish peroxidase conjugated rabbit antibody to mouse immunoglobulin (Dako, Glostrup, Denmark). The signal was visualized with diaminobenzidine, which produces a reddish-brown precipitate. Blue counterstining was performed with haematoxyline.

EXAMPLE 2

This example demonstrates the interaction of Tcf-4 and β-catenin and their function as a transcriptional activating factor.

To investigate whether hTcf-4 functionally interacts with β-catenin, we used two sets of reporter constructs in a β-catenin-Tcf reporter gene assay (7). One contained three copies of the optimal Tcf motif CCTTTGATC, or three copies of the mutant motif CCTTTGGCC, upstream of a minimal c-Fos promoter driven-luciferase expression (PTOPFLASH and PFOPFLASH). The second set contained three copies of the optimal motif, or three copies of the mutant motif, upstream of a minimal Herpes virus thymidine kinase promoter driven-Chloramphenicol Acetyl Transferase (CAT) expression (PTOPCAT and PFOPCAT, respectively). Reporter gene assays were performed as in (7). In brief, 2×10^6 cells were transfected with plasmids by electroporation. After 24 hours, cells were harvested and lysed in 1 mM DTT, 1 % Triton X-100, 15% glycerol, 25 mM Tris pH 7.8 and 8 mM $MgCl_2$. cDNAs encoding Myc-tagged versions of β-catenin and hTcf-4 were inserted into the mammalian expression vector pCDNA (Invitrogen). PCATCONTROL, encoding the CAT enzyme under the control of the SV40 promoter, was purchased from Promega.

Epitope-tagged hTcf-4 and a deletion mutant lacking, the NH_2-terminal 30 amino acids (ΔNhTcf-4) were cloned into the expression vector pCDNA. Transient transfections were performed in a murine B cell line (IIA1.6), that does not express any of the Tcf genes (6).

The TOPFLASH reporter was strongly transcribed upon cotransfection with the combination of β-catenin and hTcf-4 plasmids, but not with the individual plasmids or with the combination of β-catenin and ΔNhTcf-4 plasmids. No enhanced transcription was detected in cells transfected with the negative control PFOPFLASH (FIG. 3A). These results show that interaction of the NH_2-terminus of hTcf-4 with β-catenin results in transcriptional activation.

EXAMPLE 3

This example demonstrates the functional regulation of CRT transcriptional activation by wild-type APC.

In three $APC^{-/-}$ carcinoma cell lines, SW480, SW620 and DLD-1 (FIG. 3B), the PTOPFLASH reporter was 5–20 fold more actively transcribed than PFOPFLASH. Importantly, transfection of SW480 cells with the reporter gene and an APC-expression vector abrogated the transcriptional activity in a dose-dependent manner (FIG. 3B). In contrast APC had no effect on a cotransfected internal control (pCATCONTROL), or on the basal transcription of PFOPFLASH (FIG. 3B). The use of PTOPCAT and PFOPCAT instead of PTOPFLASH and PFOPFLASH led to comparable observations. The constitutive transcriptional activity of Tcf reporter genes in $APC^{-/-}$ colon carcinoma cells was in

stark contrast to the inactivity of these genes in non-colonic cell lines, including IIA1.6 B cells (FIG. 3A), the C57MG breast carcinoma cell line; the Jurkat and BW5147 T cell lines; the Daudi and NS1 B cell lines; the K562 erythromyeloid cell line; the HeLa cervical carcinoma line; the HepG2 hepatoma cell line; 3T3, 3T6, and Rat-I fibroblasts; and the kidney derived SV40-transformed COS cell line (7,16).

EXAMPLE 4

This example demonstrates that a functional β-catenin-hTcf-4 complex exists constitutively in APC$^{-/-}$ cells.

We used HT29-APC$^{-/-}$ colon carcinoma cells (12), in which APC is controlled by a metallothionein promoter. Induction by Zn^{++} restores wild-type levels of APC, and leads to apoptosis (12). HT29-Gal cells which carry a Zn^{++}-inducible LacZ gene were used as a control. The only Tcf family member expressed in HT29 is hTcf-4 (FIG. 2C). In nuclear extracts from uninduced HT29 derived transfectants, we readily detected hTcf-4 by gel retardation (FIG. 4). An additional band of slightly slower mobility was also observed. The addition of a β-catenin antibody resulted in the specific retardation of the latter band, indicating that it represented a β-catenin-hTcf-4 complex (FIG. 4) (12). After Zn^{++} induction for 20 hours, the β-catenin-hTcf-4 complex was diminished sixfold relative to uncomplexed hTcf-4 in HT29-APC1, while no significant change was observed in HT29-Gal cells (FIG. 4). Importantly, the overall levels of cellular β-catenin do not change during the induction period in HT29-APC1 cells (12).

Gel retardation assays were performed as described elsewhere (7). Extracts were prepared from intact nuclei that were washed four times to avoid contamination with cytoplasmic β-catenin. As the optimal Tcf/Lef probe, we used a double-stranded 15-mer CCCTTTGATCTTACC (SEQ ID NO:3); the control probe was CCCTTTGGCCTTACC (SEQ ID NO:4). (All oligonucleotides were from Isogene, Holland). The β-catenin antibody was purchased from Transduction Laboratories Lexington, Ky.). A typical binding reaction contained 3 μg nuclear protein, 0. 1 ng radiolabeled probe, 100 ng of dIdC, in 25 μl of binding buffer (60 mm KCl, 1 mM EDTA, 1 mM DTT, 10% glycerol). Samples were incubated for 20 min at room temperature, antibody was added, and the samples incubated 20 min further.

On the basis of these data, we propose the following model. In normal colonic epithelium hTcf-4 is the only expressed member of the Tcf family. The interaction of β-catenin with hTcf-4 is regulated by APC. When appropriate extracellular signals are delivered to an epithelial cell, β-catenin accumulates in a form that is not complexed with GSK3β-APC, and that enables its nuclear transport and association with hTcf-4. The HMG domain of hTcf-4 binds in a sequence-specific fashion to the regulatory sequences of specific target genes; β-catenin supplies a transactivation domain. Thus, transcriptional activation of target genes occurs only when hTcf-4 is associated with β-catenin. The hTcf-4 target genes remain to be identified. However, the link with APC and catenin suggests that these genes may participate in the generation and turnover of epithelial cells. Upon loss of wild-type APC, monomeric β-catenin accumulates in the absence of extracellular stimuli, leading to uncontrolled transcription of the hTcf-4 target genes. The apparent de novo expression of other members of the Tcf family in some colon carcinoma cell lines might lead to a further deregulation of Tcf target gene expression by the same mechanism. The control of β-catenin -Tcf signaling is likely to be an important part of the gatekeeper function of APC (19), and its disruption an early step in malignant transformation.

EXAMPLE 5

This example demonstrates that mutant APC protein does not regulate CRT and that a complete set of 20-AA repeats in APC is required to mediate inhibition of CRT.

We tested four APC mutants (FIG. 5A) for their ability to inhibit β-catenin/Tcf-regulated transcription (CRT) in transfection assays. The first mutant, APC331Δ represents a type of mutation found in the germline of Familial Adenomatous Polyposis (FAP) patients as well as in sporadic tumors (15). The APC331Δ protein is truncated at codon 331, amino-terminal to the three 15-amino-acid (AA) β-catenin-binding repeats between codons 1020 and 1169. The second mutant, APC1309Δ, is the most common germline APC mutation (15), a 5-bp deletion that produces a frameshift at codon 1309 and truncation of the protein. The APC1309Δ protein retains the 15-AA β-catenin repeats but lacks the seven 20-AA repeats between codons 1323 and 2075 that have been implicated in binding and phosphorylation of β-catenin (18). The third mutant, APC1941Δ, represents one of the most distal somatic mutations observed in colorectal tumors (25). The APC1941Δ protein is truncated at codon 1941 and therefore contains the 15-AA repeats and all but the last two 20-AA repeats. Finally, APC2644Δ represents a germline mutation resulting from a 4-bp deletion in codon 2644. Patients with this type of unusual carboxyl-terminal mutation develop few polyps (attenuated polyposis) but have pronounced extracolonic disease, particularly desmoid tumors (26).

Each of the APC mutants was cotransfected with a CRT reporter into the SW480 colorectal cancer cell line. SW480 cells have truncated APC and constitutively active CRT which can be suppressed by exogenous WT APC. Although all four mutants produced comparable levels of APC protein after transfection, they varied in their CRT inhibitory activity. The three mutants found in patients with typical polyposis or cancer were markedly deficient in inhibition of CRT (FIG. 5B). The reduced activity of APC1309Δ and APC1941Δ suggests that β-catenin binding is not sufficient for APC-mediated inhibition of CRT and that the complete set of 20-AA repeats is required. Interestingly, the inhibitory activity of the APC2644Δ mutant associated with attenuated polyposis was comparable to that of WT APC (FIG. 5B), suggesting that the DLG-binding domain at the carboxyl-terminus of APC is not required for down-regulation of CRT.

WT and mutant APC constructs (2 μg) were transfected into 293, SW480, and HCT116 cells using Lipofectamine (GIBCO/BRL, Gaithersburg). Protein was harvested 24 hours later and subjected to immunoblot analysis with APC monoclonal antibody FE9 (23). In HCT116 and 293 cells, exogenous WT APC comigrated with the endogenous APC. In SW480 cells, APC1309Δ comigrated with the endogenous mutant APC. In all other cases, the nonfunctional APC constructs (APC331Δ, APC 1309Δ, and APC1941Δ) produced as much or more protein than the CRT-functional forms of APC (APC WT and APC 2644Δ).

EXAMPLE 6

This example demonstrates that other components of the APC-regulatory pathway are affected in some cancer cells.

We evaluated CRT in two colorectal tumor cell lines (HCT116 and SW48) that express full-length APC (FIG. 6A). Both HCT116 and SW48 displayed constitutively active CRT and, in contrast to cell lines with truncated APC

(DLD1 and SW480), this activity was not inhibited by exogenous WT APC (FIG. 5B, 6B). Other (noncolorectal cancer) cell lines expressing WT APC do not display constitutive CRT activity. These transfection results suggested that the constitutive CRT in HCT116 and SW48 might be due to an altered downstream component of the APC tumor suppressor pathway.

EXAMPLE 7

This example demonstrates a defect in the gene encoding β-catenin in some cancer cells, which affects CRT.

We evaluated the status of a likely candidate for a downstream component of the APC tumor suppressor pathway, β-catenin, in the same four lines. All four lines expressed similar amounts of apparently intact β-catenin, as assessed by immunoblots (FIG. 7A). However, sequence analysis revealed that both HCT116 and SW48 harbored mutations in the β-catenin gene (CTNNB1) (FIG. 7B). HCT116 had a 3-bp deletion that removed one AA (Ser-45), and SW48 had a C to A missense mutation that changed Ser-33 to Tyr. Analysis of paraffin-embedded archival tissue from the HCT116 patient confirmed the somatic nature of this mutation and its presence in the primary tumor prior to culture. Interestingly, both mutations affected serines that have been implicated in the downregulation of β-catenin through phosphorylation by the ZW3/GSK3β kinase in Xenopus embryos (FIG. 7C) (27,28).

Genomic DNA was isolated from paraffin-embedded normal and tumor tissue from the patient from whom the HCT116 cell line was derived. A 95 bp PCR product encompassing the mutation was then amplified by PCR and directly sequenced using THERMOSEQUENASE (Amersham). The 3 bp deletion was observed in tumor but not in normal tissue.

To test the generality of this mutational mechanism, we evaluated five primary colorectal cancers in which sequencing of the entire coding region of APC revealed no mutations (25). Three of these five tumors were found to contain CTNNB1 mutations (S45F, S45F, and T44A) that altered potential ZW3/GSK3β phosphorylation sites (FIG. 7C). Each mutation appeared to affect only one of the two CTNNB1 alleles and to be somatic.

Genomic DNA was isolated from frozen-sectioned colorectal cancers and a 1001 bp PCR product containing exon 3 of CTNNB1 was then amplified by PCR and directly sequenced using ThermoSequenase (Amersham). An ACC to GCC change at codon 41 (T41A) and a TCT to TTT at codon 45 (S45F) was observed in one and two tumors, respectively.

EXAMPLE 8

This example demonstrates dominant mutations of CTNNB1 that render CRT insensitive to the effects of WT APC.

Because the β-catenin mutations were heterozygous, we hypothesized that the mutations might exert a dominant effect, rendering a fraction of cellular β-catenin insensitive to APC-mediated down regulation. To test this notion, we performed gel shift analyses with nuclear extracts from untransfected HCT116 cells. In contrast to noncolorectal cancer cell lines with intact APC, HCT116 cells contained a β-catenin/Tcf complex that gel-shifted an optimized Tcf-binding oligonucleotide, and this complex supershifted with anti-p-catenin (FIG. 8A). We also constructed β-catenin expression vectors and compared the biologic activity of the mutant β-catenin from HCT116 (β-Cat Δ45) and SW48 (β-Cat S33Y) with that of their WT counterpart. For these experiments, we used the 293 kidney epithelial cell line as it is highly transfectable, exhibits low endogenous CRT, and contains a high level of endogenous APC (FIG. 6A). In the presence of endogenous APC, both mutant β-catenins were at least 6-fold more active than the WT protein and this activity was inhibited by dominant-negative hTcf-4 (FIG. 8B).

Together, these results indicate that disruption of APC-mediated regulation of CRT is critical for colorectal tumorigenesis. This is most commonly achieved by recessive inactivating mutations of both APC alleles but, as shown here, can also be achieved by dominant mutations of CTNNB1 that render CRT insensitive to the effects of WT APC. Our results suggest that APC inhibition of CRT requires phosphorylation of β-catenin at multiple sites. These potential phosphorylation sites are consistent with the known specificity of ZW3/β SK3P (29) a serine kinase that negatively regulates β-catenin in Xenopus and Drosophila cells (27) and that interacts with APC and β-catenin in mammalian cells (23). These results also suggest a functional basis for the occasional CTNNB1 mutations observed in other tumor types (30) and illustrate how a critical pathway in human disease can be illuminated by the discovery of mutations in different components of the pathway. The next step in understanding APC function will be the identification of the genes that are activated by hTcf-4/β-catenin complexes and inhibited by WT APC. These genes are likely to be related to APC's ability to induce apoptosis in colorectal cancer cells (31).

REFERENCES

1. B. Rubinfeld et al *Science,* 262, 1731 (1993); L. K. Su, B. Vogelstein, K. W. Kinzler, ibid 262, 1734 (1993).
2. B. Gumbiner, *Curr. Opin. Cell Biol.* 7, 634 (1995).
3. B. Rubinfeld et al, *Science* 272, 1023 (1996).
4. J. Papkoff, B. Rubinfeld, B. Schryver, P. Polakis, *Mol. Cell. BioL* 16, 2128 (1996).
5. S. Munemitsa, B. Souza, I. Albert, B. Rubinfeld, P. Polakis, Proc. *Natl. Acad Sci. U.S.A.* 92, 3046 (1995); B. Rubinfeld, B. Souza, I. Albert, S. Muneinitsa, P. Polakis, *J Biol Chem.* 270, 5549 (1995).
6. M. Molenaar et al, *Cell* 86, 396 (1996); J. Behrens et al, *Nature* 382, 638 (1996); O. Huber et al., *Mech. Dev.* 59, 3 (1996).
7. M. van de Weterinc,, M. Oosterwegel, D. Dooijes, H. Clevers, *EMBO J* 10, 123 (1991); M. van de Wetering, J. Castrop, V. Korinek, Mol Cell Biol, 16, 745 (1996).
8. A. Travis et al. *Genes Dev.* 5, 880 (199 1); M. L. Waterman, W. H. Fischer, K. A. Jones ibid. p. 6562. H. Clevers and R. Grosschedl, *Immunol. Today* 17, 336 (1996).
9. J. Castrop, K. van Norren, H. C. Clevers. *Nucleic Acids Res.* 20, 611 (1992).
10. E. van Hoffen et al, *Am. J Pathol* 149, 1991 (1996).
11. M. van de Wetering, M. Oosterwegel, K. van Norren, H. Clevers, *EMBO J.* 12, 3847 (1993)
12. P. Morin, B. Vogelstein, K. W. Kinzler, *Proc. Natl. Acad. Sci. U.S.A.* 93, 7950 (1996).
13. K. W. Kinzler and B. Vogelstein, *Cell* 87, 159 (1996).
14. About 50% of the Western population develop colorectal adenomas by the age of 70 [D. Ransohoff and C. Lang, *N. Eng.l J. Med.* 325, 37 (1991)] and at least

85% of these tumors contain APC mutations; Y. Miyoshi et al., *Hum Mol Genet* 1, 229–33 (1992); J. Jen et al., *Cancer Res.* 54, 5523 (1994).

15. H. Nagase and Y. Nakamura, *Hum. Mutation* 2, 425 (1993).
16. K. W. Kinzler and B. Vogelstein, *Cell* 87, 159 (1996); S. M. Prescott and R. L. White, ibid., p. 783.
17. G. Joslyn, D. S. Richardson, R. White, T. Alber, *Proc. Natl. Acad. Sci. U.S.A.* 90, 11109 (1993); L. K. Su et al., *Cancer Res.* 53, 2728 (1993).
18. B. Rubinfeld et al., *Science* 262, 1731 (1993); L. K. Su, B. Vogelstein, K. W. Kinzler, ibid., p. 1734.
19. J. Hulsken, J. Behrens, W. Birchmeier, *Curr. Opin. Cell. Biol.* 6, 711 (1994); B. Rubinfeld, B. Souza, I. Albert, S. Munemitsu, P. Polakis, *J. Biol. Chem.* 270, 5549 (1995).
20. S. Munemitsu et al., *Cancer Res.* 54, 3676 (1994); K. J. Smith et al., ibid. p. 3672.
21. L. K. Su et al., *Cancer Res.* 55, 2972 (1995).
22. A. Matsumine et al., *Science* 272, 1020 (1996).
23. B. Rubinfeld et al., *Science* 272, 1023 (1996).
24. M. Molenaar et al., *Cell* 86, 391 (1996); J. Behrens et al., *Nature* 382, 638 (1996).
25. S. M. Powell et al., *Nature* 359, 235 (1992).
26. D. M. Eccles et al., *Am. J. of Hum. Genet.* 59, 1193 (1996); W. Friedl et al., *Hum Genet* 97, 579 (1996); R. J. Scott et al., *Human Molecular Genetics* 5, 1921 (1996).
27. C. Yost et al., *Genes Dev.* 10, 1443 (1996).
28. S. Munemitsu, I. Albert, B. Rubinfeld, P. Polakis, *Mol Cell Biol* 16, 4088 (1996).
29. M. Peifer, L. M. Pai, M. Casey, *Dev. Biol.* 166, 543 (1994).
30. D. J. Kawanishi, et al., *Mol. Cell Biol.* 15, 1175 (1995); P. F. Robbins, et al., *J. Exp. Med.* 183, 1185 (1996).
31. P. J. Morin, B. Vogelstein, K. W. Kinzler, *Proc. Natl. Acad. Sci. U.S.A.* 93, 7950 (1996).
32. J. Groden et al., *Cell* 66, 589 (1991); G. Joslyn et al., ibid., p. 601; K. W. Kinzler et al, *Science* 253, 661 (1991); I. Nishisho et al., ibid., p. 665.
33. M. Peifer, S. Berg, A. B. Reynolds, *Cell* 76, 789 (1994).
34. K. J. Smith et al, *Proc. Natl. Acad. Sci. U.S.A.* 90, 2846 (1993).
35. S. E. Goelz, S. R. Hamilton, B. Vogelstein, *Biochem. Biophys. Res. Commun.* 130, 118 (1985)

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i i i) NUMBER OF SEQUENCES: 11

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2040 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
ATGCCGCAGC TGAACGGCGG TGGAGGGGAT GACCTAGGCG CCAACGACGA ACTGATTTCC     60
TTCAAAGACG AGGGCGAACA GGAGGAGAAG AGCTCCGAAA ACTCCTCGGC AGAGAGGGAT    120
TTAGCTGATG TCAAATCGTC TCTAGTCAAT GAATCAGAAA CGAATCAAAA CAGCTCCTCC    180
GATTCCGAGG CGGAAAGACG GCCTCCGCCT CGCTCCGAAA GTTTCCGAGA CAAATCCCGG    240
GAAAGTTTGG AAGAAGCGGC CAAGAGGCAA GATGGAGGGC TCTTTAAGGG GCCACCGTAT    300
CCCGGCTACC CCTTCATCAT GATCCCCGAC CTGACGAGCC CCTACCTCCC CAAGCGATCC    360
GTCTCGCCCA CCGCCCGAAC CTATCTCCAG ATGAAATGGC CACTGCTTGA TGTCCAGGCA    420
GGGAGCCTCC AGAGTAGACA AGCCCTCAAG GATGCCCGGT CCCCATCACC GGCACACATT    480
GTCTCTAACA AAGTGCCAGT GGTGCAGCAC CCTCACCATG TCCACCCCCT CACGCCTCTT    540
ATCACGTACA GCAATGAACA CTTCACGCCG GGAAACCCAC CTCCACACTT ACCAGCCGAC    600
GTAGACCCCA AAACAGGAAT CCCACGGCCT CCGCACCCTC CAGATATATC CCCGTATTAC    660
CCACTATCGC CTGGCACCGT AGGACAAATC CCCCATCCGC TAGGATGGTT AGTACCACAG    720
CAAGGTCAAC CAGTGTACCC AATCACGACA GGAGGATTCA GACACCCCTA CCCCACAGCT    780
CTGACCGTCA ATGCTTCCGT GTCCAGGTTC CCTCCCCATA TGGTCCCACC ACATCATACG    840
```

-continued

```
CTACACACGA CGGGCATTCC GCATCCGGCC ATAGTCACAC CAACAGTCAA ACAGGAATCG   900
TCCCAGAGTG ATGTCGGCTC ACTCCATAGT TCAAAGCATC AGGACTCCAA AAAGGAAGAA   960
GAAAAGAAGA AGCCCCACAT AAAGAAACCT CTTAATGCAT TCATGTTGTA TATGAAGGAA  1020
ATGAGAGCAA AGGTCGTAGC TGAGTGCACG TTGAAAGAAA GCGCGGCCAT CAACCAGATC  1080
CTTGGGCGGA GGTGGCATGC ACTGTCCAGA GAAGAGCAAG CGAAATACTA CGAGCTGGCC  1140
CGGAAGGAGC GACAGCTTCA TATGCAACTG TACCCCGGCT GGTCCGCGCG GGATAACTAT  1200
GGAAAGAAGA AGAAGAGGAA AAGGGACAAG CAGCCGGGAG AGACCAATGG AGAAAAAAAA  1260
AGTGCGTTCG CTACATACAA GGTGAAGGCA GCTGCCTCAG CCCACCCTCT TCAGATGGAA  1320
GCTTACTAGA TTCGCCTCCC CCCTCCCCGA ACCTGCTAGG CTCCCCTCCC CGAGACGCCA  1380
AGTCACAGAC TGAGCAGACC CAGCCTCTGT CGCTGTCCCT GAAGCCCGAC CCCCTGGCCC  1440
ACCTGTCCAT GATGCCTCCG CCACCCGCCC TCCTGCTCGC TGAGGCCACC CACAAGGCCT  1500
CCGCCCTCTG TCCCAACGGG GCCCTGGACC TGCCCCCAGC CGCTTTGCAG CCTGCCGCCC  1560
CCTCCTCATC AATTGCACAG CCGTCGACTT CTTGGTTACA TTCCCACAGC TCCCTGGCCG  1620
GGACCCAGCC CCAGCCGCTG TCGCTCGTCA CCAAGTCTTT AGAATAGCTT TAGCGTCGTG  1680
AACCCCGCTG CTTTGTTTAT GGTTTTGTTT CACTTTTCTT AATTTGCCCC CCACCCCCAC  1740
CTTGAAAGGT TTTGTTTTGT ACTCTCTTAA TTTTGTGCCA TGTGGCTACA TTAGTTGATG  1800
TTTATCGAGT TCATTGGTCA ATATTTGACC CATTCTTATT TCAATTTCTC CTTTTAAATA  1860
TGTAGATGAG AGAAGAACCT CATGATTGGT ACCAAAATTT TTATCAACAG CTGTTTAAAG  1920
TCTTTGTAGC GTTTAAAAAA TATATATATA TACATAACTG TTATGTAGTT CGGATAGCTT  1980
AGTTTTAAAA GACTGATTAA AAAACAAAAA AAAAAAAAGC TTGCGAGGGA TCCCCCGGGA  2040
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2444 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
GGTTTTTTTT TTTTACCCCC CTTTTTTATT TATTATTTTT TTGCACATTG AGCGGATCCT    60
TGGGAACGAG AGAAAAAAGA AACCCAAACT CACGCGTGCA GAAGATCTCC CCCCCCTTCC   120
CCTCCCCTCC TCCCTCTTTT CCCCTCCCCA GGAGAAAAAG ACCCCCAAGC AGAAAAAAGT   180
TCACCTTGGA CTCGTCTTTT TCTTGCAATA TTTTTTGGGG GGGCAAAACT TTGAGGGGGT   240
GATTTTTTTT GGCTTTTCTT CCTCCTTCAT TTTTCTTCCA AAATTGCTGC TGGTGGGTGA   300
AAAAAAAATG CCGCAGCTGA ACGGCGGTGG AGGGGATGAC CTAGGCGCCA ACGACGAACT   360
GATTTCCTTC AAAGACGAGG GCGAACAGGA GGAGAAGAGC TCCGAAAACT CCTCGGCAGA   420
GAGGGATTTA GCTGATGTCA AATCGTCTCT AGTCAATGAA TCAGAAACGA ATCAAAACAG   480
CTCCTCCGAT TCCGAGGCGG AAAGACGGCC TCCGCCTCGC TCCGAAAGTT TCCGAGACAA   540
ATCCCGGGAA AGTTTGGAAG AAGCGGCCAA GAGGCAAGAT GGAGGGCTCT TTAAGGGGCC   600
ACCGTATCCC GGCTACCCCT TCATCATGAT CCCCGACCTG ACGAGCCCCT ACCTCCCCAA   660
GCGATCCGTC TCGCCCACCG CCCGAACCTA TCTCCAGATG AAATGGCCAC TGCTTGATGT   720
CCAGGCAGGG AGCCTCCAGA GTAGACAAGC CCTCAAGGAT GCCCGGTCCC CATCACCGGC   780
ACACATTGTC TCTAACAAAG TGCCAGTGGT GCAGCACCCT CACCATGTCC ACCCCCTCAC   840
GCCTCTTATC ACGTACAGCA ATGAACACTT CACGCCGGGA AACCCACCTC CACACTTACC   900
```

-continued

AGCCGACGTA GACCCCAAAA CAGGAATCCC ACGGCCTCCG CACCCTCCAG ATATATCCCC 960

GTATTACCCA CTATCGCCTG GCACCGTAGG ACAAATCCCC CATCCGCTAG GATGGTTAGT 1020

ACCACAGCAA GGTCAACCAG TGTACCCAAT CACGACAGGA GGATTCAGAC ACCCCTACCC 1080

CACAGCTCTG ACCGTCAATG CTTCCGTGTC CAGGTTCCCT CCCCATATGG TCCCACCACA 1140

TCATACGCTA CACACGACGG GCATTCCGCA TCCGGCCATA GTCACACCAA CAGTCAAACA 1200

GGAATCGTCC CAGAGTGATG TCGGCTCACT CCATAGTTCA AAGCATCAGG ACTCCAAAAA 1260

GGAAGAAGAA AAGAAGAAGC CCCACATAAA GAAACCTCTT AATGCATTCA TGTTGTATAT 1320

GAAGGAAATG AGAGCAAAGG TCGTAGCTGA GTGCACGTTG AAAGAAAGCG CGGCCATCAA 1380

CCAGATCCTT GGGCGGAGGT GGCATGCACT GTCCAGAGAA GAGCAAGCGA AATACTACGA 1440

GCTGGCCCGG AAGGAGCGAC AGCTTCATAT GCAACTGTAC CCCGGCTGGT CCGCGCGGGA 1500

TAACTATGGA AAGAAGAAGA AGAGGAAAAG GGACAAGCAG CCGGGAGAGA CCAATGAACA 1560

CAGCGAATGT TTCCTAAATC CTTGCCTTTC ACTTCCTCCG ATTACAGACC TCAGCGCTCC 1620

TAAGAAATGC CGAGCGCGCT TTGGCCTTGA TCAACAGAAT AACTGGTGCG GCCCTTGCAG 1680

GAGAAAAAAA AAGTGCGTTC GCTACATACA AGGTGAAGGC AGCTGCCTCA GCCCACCCTC 1740

TTCAGATGGA AGCTTACTAG ATTCGCCTCC CCCCTCCCCG AACCTGCTAG GCTCCCCTCC 1800

CCGAGACGCC AAGTCACAGA CTGAGCAGAC CCAGCCTCTG TCGCTGTCCC TGAAGCCCGA 1860

CCCCCTGGCC CACCTGTCCA TGATGCCTCC GCCACCCGCC CTCCTGCTCG CTGAGGCCAC 1920

CCACAAGGCC TCCGCCCTCT GTCCCAACGG GGCCCTGGAC CTGCCCCCAG CCGCTTTGCA 1980

GCCTGCCGCC CCCTCCTCAT CAATTGCACA GCCGTCGACT TCTTGGTTAC ATTCCCACAG 2040

CTCCCTGGCC GGGACCCAGC CCCAGCCGCT GTCGCTCGTC ACCAAGTCTT TAGAATAGCT 2100

TTAGCGTCGT GAACCCCGCT GCTTTGTTTA TGGTTTTGTT TCACTTTTCT TAATTTGCCC 2160

CCCACCCCCA CCTTGAAAGG TTTTGTTTTG TACTCTCTTA ATTTTGTGCC ATGTGGCTAC 2220

ATTAGTTGAT GTTTATCGAG TTCATTGGTC AATATTTGAC CCATTCTTAT TTCAATTTCT 2280

CCTTTTAAAT ATGTAGATGA GAGAAGAACC TCATGATTGG TACCAAAATT TTTATCAACA 2340

GCTGTTTAAA GTCTTTGTAG CGTTTAAAAA ATATATATAT ATACATAACT GTTATGTAGT 2400

TCGGATAGCT TAGTTTTAAA AGACTGATTA AAAAACAAAA AAAA 2444

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CCCTTTGATC TTACC 15

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: cDNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CCCTTTGGCC TTACC 15

-continued

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 442 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: None

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Pro Gln Leu Asn Gly Gly Gly Gly Asp Asp Leu Gly Ala Asn Asp
1 5 10 15

Glu Leu Ile Ser Phe Lys Asp Glu Gly Glu Gln Glu Glu Lys Ser Ser
20 25 30

Gln Asn Ser Ser Ala Glu Arg Asp Leu Ala Asp Val Lys Ser Ser Leu
35 40 45

Val Asn Glu Ser Glu Thr Asn Gln Asn Ser Ser Ser Asp Ser Glu Ala
50 55 60

Glu Arg Arg Pro Pro Pro Arg Ser Glu Ser Phe Arg Asp Lys Ser Arg
65 70 75 80

Glu Ser Leu Glu Glu Ala Ala Lys Arg Gln Asp Gly Gly Leu Phe Lys
85 90 95

Gly Pro Pro Tyr Pro Gly Tyr Pro Phe Ile Met Ile Pro Asp Leu Thr
100 105 110

Ser Pro Tyr Leu Pro Lys Arg Ser Val Ser Pro Thr Ala Arg Thr Tyr
115 120 125

Leu Gln Met Lys Trp Pro Leu Leu Asp Val Gln Ala Gly Ser Leu Gln
130 135 140

Ser Arg Gln Ala Leu Lys Asp Ala Arg Ser Pro Ser Pro Ala His Ile
145 150 155 160

Val Ser Asn Lys Val Pro Val Val Gln His Pro His His Val His Pro
165 170 175

Leu Thr Pro Leu Ile Thr Tyr Ser Asn Glu His Phe Thr Pro Gly Asn
180 185 190

Pro Pro Pro His Leu Pro Ala Asp Val Asp Pro Lys Thr Gly Ile Pro
195 200 205

Arg Pro Pro His Pro Pro Asp Ile Ser Pro Tyr Tyr Pro Leu Ser Pro
210 215 220

Gly Thr Val Gly Gln Ile Pro His Pro Leu Gly Trp Leu Val Pro Gln
225 230 235 240

Gln Gly Gln Pro Val Tyr Pro Ile Thr Thr Gly Gly Phe Arg His Pro
245 250 255

Tyr Pro Thr Ala Leu Thr Val Asn Ala Ser Val Ser Arg Phe Pro Pro
260 265 270

His Met Val Pro Pro His His Thr Leu His Thr Thr Gly Ile Pro His
275 280 285

Pro Ala Ile Val Thr Pro Thr Val Lys Gln Glu Ser Ser Gln Ser Asp
290 295 300

Val Gly Ser Leu His Ser Ser Lys His Gln Asp Ser Lys Lys Glu Glu
305 310 315 320

Glu Lys Lys Lys Pro His Ile Lys Lys Pro Leu Asn Ala Phe Met Leu
325 330 335

Tyr Met Lys Glu Met Arg Ala Lys Val Val Ala Glu Cys Thr Leu Lys
340 345 350

-continued

```
Glu Ser Ala Ala Ile Asn Gln Ile Leu Gly Arg Arg Trp His Ala Leu
            355                 360                 365

Ser Arg Glu Glu Gln Ala Lys Tyr Tyr Glu Leu Ala Arg Lys Glu Arg
        370                 375                 380

Gln Leu His Met Gln Leu Tyr Pro Gly Trp Ser Ala Arg Asp Asn Tyr
385                 390                 395                 400

Gly Lys Lys Lys Lys Arg Lys Arg Asp Lys Gln Pro Gly Glu Thr Asn
                405                 410                 415

Gly Glu Lys Lys Ser Ala Phe Ala Thr Tyr Lys Val Lys Ala Ala Ala
            420                 425                 430

Ser Ala His Pro Leu Gln Met Glu Ala Tyr
        435                 440
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 596 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: None

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
Met Pro Gln Leu Asn Gly Gly Gly Gly Asp Asp Leu Gly Ala Asn Asp
 1               5                  10                  15

Glu Leu Ile Ser Phe Lys Asp Glu Gly Glu Gln Glu Glu Lys Ser Ser
            20                  25                  30

Glu Asn Ser Ser Ala Glu Arg Asp Leu Ala Asp Val Lys Ser Ser Leu
        35                  40                  45

Val Asn Glu Ser Glu Thr Asn Gln Asn Ser Ser Ser Asp Ser Glu Ala
    50                  55                  60

Glu Arg Arg Pro Pro Pro Arg Ser Glu Ser Phe Arg Asp Lys Ser Arg
65                  70                  75                  80

Glu Ser Leu Glu Glu Ala Ala Lys Arg Gln Asp Gly Gly Leu Phe Lys
                85                  90                  95

Gly Pro Pro Tyr Pro Gly Tyr Pro Phe Ile Met Ile Pro Asp Leu Thr
            100                 105                 110

Ser Pro Tyr Leu Pro Asn Gly Ser Val Ser Pro Thr Ala Arg Thr Tyr
        115                 120                 125

Leu Gln Met Lys Trp Pro Leu Leu Asp Val Gln Ala Gly Ser Leu Gln
    130                 135                 140

Ser Arg Gln Ala Leu Lys Asp Ala Arg Ser Pro Ser Pro Ala His Ile
145                 150                 155                 160

Val Ser Asn Lys Val Pro Val Val Gln His Pro His His Val His Pro
                165                 170                 175

Leu Thr Pro Leu Ile Thr Tyr Ser Asn Glu His Phe Thr Pro Gly Asn
            180                 185                 190

Pro Pro Pro His Leu Pro Ala Asp Val Asp Pro Lys Thr Gly Ile Pro
        195                 200                 205

Arg Pro Pro His Pro Pro Asp Ile Ser Pro Tyr Tyr Pro Leu Ser Pro
    210                 215                 220

Gly Thr Val Gly Gln Ile Pro His Pro Leu Gly Trp Leu Val Pro Gln
225                 230                 235                 240

Gln Gly Gln Pro Val Tyr Pro Ile Thr Thr Gly Gly Phe Arg His Pro
                245                 250                 255

Tyr Pro Thr Ala Leu Thr Val Asn Ala Ser Val Ser Arg Phe Pro Pro
```

-continued

260 265 270

His Met Val Pro Pro His His Thr Leu His Thr Thr Gly Ile Pro His
275 280 285

Pro Ala Ile Val Thr Pro Thr Val Lys Gln Glu Ser Ser Gln Ser Asp
290 295 300

Val Gly Ser Leu His Ser Ser Lys His Gln Asp Ser Lys Lys Glu Glu
305 310 315 320

Glu Lys Lys Lys Pro His Ile Lys Lys Pro Leu Asn Ala Phe Met Leu
325 330 335

Tyr Met Lys Glu Met Arg Ala Lys Val Val Ala Glu Cys Thr Leu Lys
340 345 350

Glu Ser Ala Ala Ile Asn Gln Ile Leu Gly Arg Arg Trp His Ala Leu
355 360 365

Ser Arg Glu Glu Gln Ala Lys Tyr Tyr Glu Leu Ala Arg Lys Glu Arg
370 375 380

Gln Leu His Met Gln Leu Tyr Pro Gly Trp Ser Ala Arg Asp Asn Tyr
385 390 395 400

Gly Lys Lys Lys Lys Arg Lys Arg Asp Lys Gln Pro Gly Glu Thr Asn
405 410 415

Glu His Ser Glu Cys Phe Leu Asn Pro Cys Leu Ser Leu Pro Pro Ile
420 425 430

Thr Asp Leu Ser Ala Pro Lys Lys Cys Arg Ala Arg Phe Gly Leu Asp
435 440 445

Gln Gln Asn Asn Trp Cys Gly Pro Cys Arg Arg Lys Lys Lys Cys Val
450 455 460

Arg Tyr Ile Gln Gly Glu Gly Ser Cys Leu Ser Pro Pro Ser Ser Asp
465 470 475 480

Gly Ser Leu Leu Asp Ser Pro Pro Pro Ser Pro Asn Leu Leu Gly Ser
485 490 495

Pro Pro Arg Asp Ala Lys Ser Gln Thr Glu Gln Thr Gln Pro Leu Ser
500 505 510

Leu Ser Leu Lys Pro Asp Pro Leu Ala His Leu Ser Met Met Pro Pro
515 520 525

Pro Pro Ala Leu Leu Leu Ala Glu Ala Thr His Lys Ala Ser Ala Leu
530 535 540

Cys Pro Asn Gly Ala Leu Asp Leu Pro Pro Ala Ala Leu Gln Pro Ala
545 550 555 560

Ala Pro Ser Ser Ser Ile Ala Gln Pro Ser Thr Ser Trp Leu His Ser
565 570 575

His Ser Ser Leu Ala Gly Thr Gln Pro Gln Pro Leu Ser Leu Val Thr
580 585 590

Lys Ser Leu Glu
595

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2973 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: None

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Ala Ala Ala Ser Tyr Asp Gln Leu Leu Lys Gln Val Glu Ala Leu
1 5 10 15

-continued

Lys Met Glu Asn Ser Asn Leu Arg Gln Glu Leu Glu Asp Asn Ser Asn
20 25 30

His Leu Thr Lys Leu Glu Thr Glu Ala Ser Asn Met Lys Glu Val Leu
35 40 45

Lys Gln Leu Gln Gly Ser Ile Glu Asp Glu Ala Met Ala Ser Ser Gly
50 55 60

Gln Ile Asp Leu Leu Glu Arg Leu Lys Glu Leu Asn Leu Asp Ser Ser
65 70 75 80

Asn Phe Pro Gly Val Lys Leu Arg Ser Lys Met Ser Leu Arg Ser Tyr
85 90 95

Gly Ser Arg Glu Gly Ser Val Ser Ser Arg Ser Gly Glu Cys Ser Pro
100 105 110

Val Pro Met Gly Ser Phe Pro Arg Arg Gly Phe Val Asn Gly Ser Arg
115 120 125

Glu Ser Thr Gly Tyr Leu Glu Glu Leu Glu Lys Glu Arg Ser Leu Leu
130 135 140

Leu Ala Asp Leu Asp Lys Glu Glu Lys Glu Lys Asp Trp Tyr Tyr Ala
145 150 155 160

Gln Leu Gln Asn Leu Thr Lys Arg Ile Asp Ser Leu Pro Leu Thr Glu
165 170 175

Asn Phe Ser Leu Gln Thr Asp Met Thr Arg Arg Gln Leu Glu Tyr Glu
180 185 190

Ala Arg Gln Ile Arg Val Ala Met Glu Glu Gln Leu Gly Thr Cys Gln
195 200 205

Asp Met Glu Lys Arg Ala Gln Arg Arg Ile Ala Arg Ile Gln Gln Ile
210 215 220

Glu Lys Asp Ile Leu Arg Ile Arg Gln Leu Leu Gln Ser Gln Ala Thr
225 230 235 240

Glu Ala Glu Arg Ser Ser Gln Asn Lys His Glu Thr Gly Ser His Asp
245 250 255

Ala Glu Arg Gln Asn Glu Gly Gln Gly Val Gly Glu Ile Asn Met Ala
260 265 270

Thr Ser Gly Asn Gly Gln Gly Ser Thr Thr Arg Met Asp His Glu Thr
275 280 285

Ala Ser Val Leu Ser Ser Ser Ser Thr His Ser Ala Pro Arg Arg Leu
290 295 300

Thr Ser His Leu Gly Thr Lys Val Glu Met Val Tyr Ser Leu Leu Ser
305 310 315 320

Met Leu Gly Thr His Asp Lys Asp Asp Met Ser Arg Thr Leu Leu Ala
325 330 335

Met Ser Ser Ser Gln Asp Ser Cys Ile Ser Met Arg Gln Ser Gly Cys
340 345 350

Leu Pro Leu Leu Ile Gln Leu Leu His Gly Asn Asp Lys Asp Ser Val
355 360 365

Leu Leu Gly Asn Ser Arg Gly Ser Lys Glu Ala Arg Ala Arg Ala Ser
370 375 380

Ala Ala Leu His Asn Ile Ile His Ser Gln Pro Asp Asp Lys Arg Gly
385 390 395 400

Arg Arg Glu Ile Arg Val Leu His Leu Leu Glu Gln Ile Arg Ala Tyr
405 410 415

Cys Glu Thr Cys Trp Glu Trp Gln Glu Ala His Glu Pro Gly Met Asp
420 425 430

Gln Asp Lys Asn Pro Met Pro Ala Pro Val Glu His Gln Ile Cys Pro

-continued

435 440 445

Ala Val Cys Val Leu Met Lys Leu Ser Phe Asp Glu Glu His Arg His
450 455 460

Ala Met Asn Glu Leu Gly Gly Leu Gln Ala Ile Ala Glu Leu Leu Gln
465 470 475 480

Val Asp Cys Glu Met Tyr Gly Leu Thr Asn Asp His Tyr Ser Ile Thr
485 490 495

Leu Arg Arg Tyr Ala Gly Met Ala Leu Thr Asn Leu Thr Phe Gly Asp
500 505 510

Val Ala Asn Lys Ala Thr Leu Cys Ser Met Lys Gly Cys Met Arg Ala
515 520 525

Leu Val Ala Gln Leu Lys Ser Glu Ser Glu Asp Leu Gln Gln Val Ile
530 535 540

Ala Ser Val Leu Arg Asn Leu Ser Trp Arg Ala Asp Val Asn Ser Lys
545 550 555 560

Lys Thr Leu Arg Glu Val Gly Ser Val Lys Ala Leu Met Glu Cys Ala
565 570 575

Leu Glu Val Lys Lys Glu Ser Thr Leu Lys Ser Val Leu Ser Ala Leu
580 585 590

Trp Asn Leu Ser Ala His Cys Thr Glu Asn Lys Ala Asp Ile Cys Ala
595 600 605

Val Asp Gly Ala Leu Ala Phe Leu Val Gly Thr Leu Thr Tyr Arg Ser
610 615 620

Gln Thr Asn Thr Leu Ala Ile Ile Glu Ser Gly Gly Gly Ile Leu Arg
625 630 635 640

Asn Val Ser Ser Leu Ile Ala Thr Asn Glu Asp His Arg Gln Ile Leu
645 650 655

Arg Glu Asn Asn Cys Leu Gln Thr Leu Leu Gln His Leu Lys Ser His
660 665 670

Ser Leu Thr Ile Val Ser Asn Ala Cys Gly Thr Leu Trp Asn Leu Ser
675 680 685

Ala Arg Asn Pro Lys Asp Gln Glu Ala Leu Trp Asp Met Gly Ala Val
690 695 700

Ser Met Leu Lys Asn Leu Ile His Ser Lys His Lys Met Ile Ala Met
705 710 715 720

Gly Ser Ala Ala Ala Leu Arg Asn Leu Met Ala Asn Arg Pro Ala Lys
725 730 735

Tyr Lys Asp Ala Asn Ile Met Ser Pro Gly Ser Ser Leu Pro Ser Leu
740 745 750

His Val Arg Lys Gln Lys Ala Leu Glu Ala Glu Leu Asp Ala Gln His
755 760 765

Leu Ser Glu Thr Phe Asp Asn Ile Asp Asn Leu Ser Pro Lys Ala Ser
770 775 780

His Arg Ser Lys Gln Arg His Lys Gln Ser Leu Tyr Gly Asp Tyr Val
785 790 795 800

Phe Asp Thr Asn Arg His Asp Asp Asn Arg Ser Asp Asn Phe Asn Thr
805 810 815

Gly Asn Met Thr Val Leu Ser Pro Tyr Leu Asn Thr Thr Val Leu Pro
820 825 830

Ser Ser Ser Ser Ser Arg Gly Ser Leu Asp Ser Ser Arg Ser Glu Lys
835 840 845

Asp Arg Ser Leu Glu Arg Glu Arg Gly Ile Gly Leu Gly Asn Tyr His
850 855 860

-continued

Pro Ala Thr Glu Asn Pro Gly Thr Ser Ser Lys Arg Gly Leu Gln Ile
865 870 875 880

Ser Thr Thr Ala Ala Gln Ile Ala Lys Val Met Glu Glu Val Ser Ala
885 890 895

Ile His Thr Ser Gln Glu Asp Arg Ser Ser Gly Ser Thr Thr Glu Leu
900 905 910

His Cys Val Thr Asp Gln Arg Asn Ala Leu Arg Arg Ser Ser Ala Ala
915 920 925

His Thr His Ser Asn Thr Tyr Asn Phe Thr Lys Ser Glu Asn Ser Asn
930 935 940

Arg Thr Cys Ser Met Pro Tyr Ala Lys Leu Glu Tyr Lys Arg Ser Ser
945 950 955 960

Asn Asp Ser Leu Asn Ser Val Ser Ser Ser Asp Gly Tyr Gly Lys Arg
965 970 975

Gly Gln Met Lys Pro Ser Ile Glu Ser Tyr Ser Glu Asp Asp Glu Ser
980 985 990

Lys Phe Cys Ser Tyr Gly Gln Tyr Pro Ala Asp Leu Ala His Lys Ile
995 1000 1005

His Ser Ala Asn His Met Asp Asp Asn Asp Gly Glu Leu Asp Thr Pro
1010 1015 1020

Ile Asn Tyr Ser Leu Lys Tyr Ser Asp Glu Gln Leu Asn Ser Gly Arg
1025 1030 1035 1040

Gln Ser Pro Ser Gln Asn Glu Arg Trp Ala Arg Pro Lys His Ile Ile
1045 1050 1055

Glu Asp Glu Ile Lys Gln Ser Glu Gln Arg Gln Ser Arg Asn Gln Ser
1060 1065 1070

Thr Thr Tyr Pro Val Tyr Thr Glu Ser Thr Asp Asp Lys His Leu Lys
1075 1080 1085

Phe Gln Pro His Phe Gly Gln Gln Glu Cys Val Ser Pro Tyr Arg Ser
1090 1095 1100

Arg Gly Ala Asn Gly Ser Glu Thr Asn Arg Val Gly Ser Asn His Gly
1105 1110 1115 1120

Ile Asn Gln Asn Val Ser Gln Ser Leu Cys Gln Glu Asp Asp Tyr Glu
1125 1130 1135

Asp Asp Lys Pro Thr Asn Tyr Ser Glu Arg Tyr Ser Glu Glu Glu Gln
1140 1145 1150

His Glu Glu Glu Glu Arg Pro Thr Asn Tyr Ser Ile Lys Tyr Asn Glu
1155 1160 1165

Glu Lys Arg His Val Asp Gln Pro Ile Asp Tyr Ser Leu Lys Tyr Ala
1170 1175 1180

Thr Asp Ile Pro Ser Ser Gln Lys Gln Ser Phe Ser Phe Ser Lys Ser
1185 1190 1195 1200

Ser Ser Gly Gln Ser Ser Lys Thr Glu His Met Ser Ser Ser Ser Glu
1205 1210 1215

Asn Thr Ser Thr Pro Ser Ser Asn Ala Lys Arg Gln Asn Gln Leu His
1220 1225 1230

Pro Ser Ser Ala Gln Ser Arg Ser Gly Gln Pro Gln Lys Ala Ala Thr
1235 1240 1245

Cys Lys Val Ser Ser Ile Asn Gln Glu Thr Ile Gln Thr Tyr Cys Val
1250 1255 1260

Glu Asp Thr Pro Ile Cys Phe Ser Arg Cys Ser Ser Leu Ser Ser Leu
1265 1270 1275 1280

Ser Ser Ala Glu Asp Glu Ile Gly Cys Asn Gln Thr Thr Gln Glu Ala
1285 1290 1295

-continued

```
Asp Ser Ala Asn Thr Leu Gln Ile Ala Glu Ile Lys Glu Lys Ile Gly
            1300                1305                1310

Thr Arg Ser Ala Glu Asp Pro Val Ser Glu Val Pro Ala Val Ser Gln
        1315                1320                1325

His Pro Arg Thr Lys Ser Ser Arg Leu Gln Gly Ser Ser Leu Ser Ser
    1330                1335                1340

Glu Ser Ala Arg His Lys Ala Val Glu Phe Ser Ser Gly Ala Lys Ser
1345                1350                1355                1360

Pro Ser Lys Ser Gly Ala Gln Thr Pro Lys Ser Pro Pro Glu His Tyr
                1365                1370                1375

Val Gln Glu Thr Pro Leu Met Phe Ser Arg Cys Thr Ser Val Ser Ser
            1380                1385                1390

Leu Asp Ser Phe Glu Ser Arg Ser Ile Ala Ser Ser Val Gln Ser Glu
        1395                1400                1405

Pro Cys Ser Gly Met Val Ser Gly Ile Ile Ser Pro Ser Asp Leu Pro
    1410                1415                1420

Asp Ser Pro Gly Gln Thr Met Pro Pro Ser Arg Ser Lys Thr Pro Pro
1425                1430                1435                1440

Pro Pro Pro Gln Thr Ala Gln Thr Lys Arg Glu Val Pro Lys Asn Lys
                1445                1450                1455

Ala Pro Thr Ala Glu Lys Arg Glu Ser Gly Pro Lys Gln Ala Ala Val
            1460                1465                1470

Asn Ala Ala Val Gln Arg Val Gln Val Leu Pro Asp Ala Asp Thr Leu
        1475                1480                1485

Leu His Phe Ala Thr Glu Ser Thr Pro Asp Gly Phe Ser Cys Ser Ser
    1490                1495                1500

Ser Leu Ser Ala Leu Ser Leu Asp Glu Pro Phe Ile Gln Lys Asp Val
1505                1510                1515                1520

Glu Leu Arg Ile Met Pro Pro Val Gln Glu Asn Asp Asn Gly Asn Glu
                1525                1530                1535

Thr Glu Ser Glu Gln Pro Lys Glu Ser Asn Glu Asn Gln Glu Lys Glu
            1540                1545                1550

Ala Glu Lys Thr Ile Asp Ser Glu Lys Asp Leu Leu Asp Asp Ser Asp
        1555                1560                1565

Asp Asp Asp Ile Glu Ile Leu Glu Glu Cys Ile Ile Ser Ala Met Pro
    1570                1575                1580

Thr Lys Ser Ser Arg Lys Ala Lys Lys Pro Ala Gln Thr Ala Ser Lys
1585                1590                1595                1600

Leu Pro Pro Pro Val Ala Arg Lys Pro Ser Gln Leu Pro Val Tyr Lys
                1605                1610                1615

Leu Leu Pro Ser Gln Asn Arg Leu Gln Pro Gln Lys His Val Ser Phe
            1620                1625                1630

Thr Pro Gly Asp Asp Met Pro Arg Val Tyr Cys Val Glu Gly Thr Pro
        1635                1640                1645

Ile Asn Phe Ser Thr Ala Thr Ser Leu Ser Asp Leu Thr Ile Glu Ser
    1650                1655                1660

Pro Pro Asn Glu Leu Ala Ala Gly Glu Gly Val Arg Gly Gly Ala Gln
1665                1670                1675                1680

Ser Gly Glu Phe Glu Lys Arg Asp Thr Ile Pro Thr Glu Gly Arg Ser
                1685                1690                1695

Thr Asp Glu Ala Gln Gly Gly Lys Thr Ser Ser Val Thr Ile Pro Glu
            1700                1705                1710

Leu Asp Asp Asn Lys Ala Glu Glu Gly Asp Ile Leu Ala Glu Cys Ile
```

-continued

1715 1720 1725

Asn Ser Ala Met Pro Lys Gly Lys Ser His Lys Pro Phe Arg Val Lys
1730 1735 1740

Lys Ile Met Asp Gln Val Gln Gln Ala Ser Ala Ser Ser Ser Ala Pro
1745 1750 1755 1760

Asn Lys Asn Gln Leu Asp Gly Lys Lys Lys Lys Pro Thr Ser Pro Val
1765 1770 1775

Lys Pro Ile Pro Gln Asn Thr Glu Tyr Arg Thr Arg Val Arg Lys Asn
1780 1785 1790

Ala Asp Ser Lys Asn Asn Leu Asn Ala Glu Arg Val Phe Ser Asp Asn
1795 1800 1805

Lys Asp Ser Lys Lys Gln Asn Leu Lys Asn Asn Ser Lys Asp Phe Asn
1810 1815 1820

Asp Lys Leu Pro Asn Asn Glu Asp Arg Val Arg Gly Ser Phe Ala Phe
1825 1830 1835 1840

Asp Ser Pro His His Tyr Thr Pro Ile Glu Gly Thr Pro Tyr Cys Phe
1845 1850 1855

Ser Arg Asn Asp Ser Leu Ser Ser Leu Asp Phe Asp Asp Asp Asp Val
1860 1865 1870

Asp Leu Ser Arg Glu Lys Ala Glu Leu Arg Lys Ala Lys Glu Asn Lys
1875 1880 1885

Glu Ser Glu Ala Lys Val Thr Ser His Thr Glu Leu Thr Ser Asn Gln
1890 1895 1900

Gln Ser Ala Asn Lys Thr Gln Ala Ile Ala Lys Gln Pro Ile Asn Arg
1905 1910 1915 1920

Gly Gln Pro Lys Pro Ile Leu Gln Lys Gln Ser Thr Phe Pro Gln Ser
1925 1930 1935

Ser Lys Asp Ile Pro Asp Arg Gly Ala Ala Thr Asp Glu Lys Leu Gln
1940 1945 1950

Asn Phe Ala Ile Glu Asn Thr Pro Val Cys Phe Ser His Asn Ser Ser
1955 1960 1965

Leu Ser Ser Leu Ser Asp Ile Asp Gln Glu Asn Asn Asn Lys Glu Asn
1970 1975 1980

Glu Pro Ile Lys Glu Thr Glu Pro Pro Asp Ser Gln Gly Glu Pro Ser
1985 1990 1995 2000

Lys Pro Gln Ala Ser Gly Tyr Ala Pro Lys Ser Phe His Val Glu Asp
2005 2010 2015

Thr Pro Val Cys Phe Ser Arg Asn Ser Ser Leu Ser Ser Leu Ser Ile
2020 2025 2030

Asp Ser Glu Asp Asp Leu Leu Gln Glu Cys Ile Ser Ser Ala Met Pro
2035 2040 2045

Lys Lys Lys Lys Pro Ser Arg Leu Lys Gly Asp Asn Glu Lys His Ser
2050 2055 2060

Pro Arg Asn Met Gly Gly Ile Leu Gly Glu Asp Leu Thr Leu Asp Leu
2065 2070 2075 2080

Lys Asp Ile Gln Arg Pro Asp Ser Glu His Gly Leu Ser Pro Asp Ser
2085 2090 2095

Glu Asn Phe Asp Trp Lys Ala Ile Gln Glu Gly Ala Asn Ser Ile Val
2100 2105 2110

Ser Ser Leu His Gln Ala Ala Ala Ala Ala Ala Cys Leu Ser Arg Gln Ala
2115 2120 2125

Ser Ser Asp Ser Asp Ser Ile Leu Ser Leu Lys Ser Gly Ile Ser Leu
2130 2135 2140

-continued

Gly Ser Pro Phe His Leu Thr Pro Asp Gln Glu Glu Lys Pro Phe Thr
2145 2150 2155 2160

Ser Asn Lys Gly Pro Arg Ile Leu Lys Pro Gly Glu Lys Ser Thr Leu
2165 2170 2175

Glu Thr Lys Lys Ile Glu Ser Glu Ser Lys Gly Ile Lys Gly Gly Lys
2180 2185 2190

Lys Val Tyr Lys Ser Leu Ile Thr Gly Lys Val Arg Ser Asn Ser Glu
2195 2200 2205

Ile Ser Gly Gln Met Lys Gln Pro Leu Gln Ala Asn Met Pro Ser Ile
2210 2215 2220

Ser Arg Gly Arg Thr Met Ile His Ile Pro Gly Val Arg Asn Ser Ser
2225 2230 2235 2240

Ser Ser Thr Ser Pro Val Ser Lys Lys Gly Pro Pro Leu Lys Thr Pro
2245 2250 2255

Ala Ser Lys Ser Pro Ser Glu Gly Gln Thr Ala Thr Thr Ser Pro Arg
2260 2265 2270

Gly Ala Lys Pro Ser Val Lys Ser Glu Leu Ser Pro Val Ala Arg Gln
2275 2280 2285

Thr Ser Gln Ile Gly Gly Ser Ser Lys Ala Pro Ser Arg Ser Gly Ser
2290 2295 2300

Arg Asp Ser Thr Pro Ser Arg Pro Ala Gln Gln Pro Leu Ser Arg Pro
2305 2310 2315 2320

Ile Gln Ser Pro Gly Arg Asn Ser Ile Ser Pro Gly Arg Asn Gly Ile
2325 2330 2335

Ser Pro Pro Asn Lys Leu Ser Gln Leu Pro Arg Thr Ser Ser Pro Ser
2340 2345 2350

Thr Ala Ser Thr Lys Ser Ser Gly Ser Gly Lys Met Ser Tyr Thr Ser
2355 2360 2365

Pro Gly Arg Gln Met Ser Gln Gln Asn Leu Thr Lys Gln Thr Gly Leu
2370 2375 2380

Ser Lys Asn Ala Ser Ser Ile Pro Arg Ser Glu Ser Ala Ser Lys Gly
2385 2390 2395 2400

Leu Asn Gln Met Asn Asn Gly Asn Gly Ala Asn Lys Lys Val Glu Leu
2405 2410 2415

Ser Arg Met Ser Ser Thr Lys Ser Ser Gly Ser Glu Ser Asp Arg Ser
2420 2425 2430

Glu Arg Pro Val Leu Val Arg Gln Ser Thr Phe Ile Lys Glu Ala Pro
2435 2440 2445

Ser Pro Thr Leu Arg Arg Lys Leu Glu Glu Ser Ala Ser Phe Glu Ser
2450 2455 2460

Leu Ser Pro Ser Ser Arg Pro Ala Ser Pro Thr Arg Ser Gln Ala Gln
2465 2470 2475 2480

Thr Pro Val Leu Ser Pro Ser Leu Pro Asp Met Ser Leu Ser Thr His
2485 2490 2495

Ser Ser Val Gln Ala Gly Gly Trp Arg Lys Leu Pro Pro Asn Leu Ser
2500 2505 2510

Pro Thr Ile Glu Tyr Asn Asp Gly Arg Pro Ala Lys Arg His Asp Ile
2515 2520 2525

Ala Arg Ser His Ser Glu Ser Pro Ser Arg Leu Pro Ile Asn Arg Ser
2530 2535 2540

Gly Thr Trp Lys Arg Glu His Ser Lys His Ser Ser Ser Leu Pro Arg
2545 2550 2555 2560

Val Ser Thr Trp Arg Arg Thr Gly Ser Ser Ser Ser Ile Leu Ser Ala
2565 2570 2575

-continued

Ser Ser Glu Ser Ser Glu Lys Ala Lys Ser Glu Asp Glu Lys His Val
2580 2585 2590

Asn Ser Ile Ser Gly Thr Lys Gln Ser Lys Glu Asn Gln Val Ser Ala
2595 2600 2605

Lys Gly Thr Trp Arg Lys Ile Lys Glu Asn Glu Phe Ser Pro Thr Asn
2610 2615 2620

Ser Thr Ser Gln Thr Val Ser Ser Gly Ala Thr Asn Gly Ala Glu Ser
2625 2630 2635 2640

Lys Thr Leu Ile Tyr Gln Met Ala Pro Ala Val Ser Lys Thr Glu Asp
2645 2650 2655

Val Trp Val Arg Ile Glu Asp Cys Pro Ile Asn Asn Pro Arg Ser Gly
2660 2665 2670

Arg Ser Pro Thr Gly Asn Thr Pro Pro Val Ile Asp Ser Val Ser Glu
2675 2680 2685

Lys Ala Asn Pro Asn Ile Lys Asp Ser Lys Asp Asn Gln Ala Lys Gln
2690 2695 2700

Asn Val Gly Asn Gly Ser Val Pro Met Arg Thr Val Gly Leu Glu Asn
2705 2710 2715 2720

Arg Leu Asn Ser Phe Ile Gln Val Asp Ala Pro Asp Gln Lys Gly Thr
2725 2730 2735

Glu Ile Lys Pro Gly Gln Asn Asn Pro Val Pro Val Ser Glu Thr Asn
2740 2745 2750

Glu Ser Ser Ile Val Glu Arg Thr Pro Phe Ser Ser Ser Ser Ser Ser
2755 2760 2765

Lys His Ser Ser Pro Ser Gly Thr Val Ala Ala Arg Val Thr Pro Phe
2770 2775 2780

Asn Tyr Asn Pro Ser Pro Arg Lys Ser Ser Ala Asp Ser Thr Ser Ala
2785 2790 2795 2800

Arg Pro Ser Gln Ile Pro Thr Pro Val Asn Asn Asn Thr Lys Lys Arg
2805 2810 2815

Asp Ser Lys Thr Asp Ser Thr Glu Ser Ser Gly Thr Gln Ser Pro Lys
2820 2825 2830

Arg His Ser Gly Ser Tyr Leu Val Thr Ser Val Lys Arg Gly Arg Met
2835 2840 2845

Lys Leu Arg Lys Phe Tyr Val Asn Tyr Asn Cys Tyr Ile Asp Ile Leu
2850 2855 2860

Phe Gln Met Lys Leu Lys Thr Glu Lys Phe Cys Lys Val Phe Leu Leu
2865 2870 2875 2880

Glu Gly Phe Cys Ser Gly Ser His Ile Tyr Thr Leu Ser Ser Leu Val
2885 2890 2895

Leu Phe Trp Glu Ala Leu Leu Met Val Arg Lys Lys Ile Val Lys Pro
2900 2905 2910

Ser Met Phe Val Gln Tyr Val Leu His Val Phe Lys Val Ala Pro Ile
2915 2920 2925

Pro Thr Ser Phe Asn Tyr Cys Leu Ser Asn Asn Glu His Tyr Arg Lys
2930 2935 2940

Ile Tyr Ile Ala Val Ile Asn His Phe Ile Ile Asn Leu Asn Leu His
2945 2950 2955 2960

Gln Gly Lys Ile Gly Ile Tyr Ala Lys Lys Asn Val Phe
2965 2970

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

-continued

(A) LENGTH: 486 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: None

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```
Met Pro Gln Leu Asp Ser Gly Gly Gly Gly Ala Gly Gly Gly Asp Asp
 1               5                  10                  15

Leu Gly Ala Pro Asp Glu Leu Leu Ala Phe Gln Asp Glu Gly Glu Glu
            20                  25                  30

Gln Asp Asp Lys Ser Arg Asp Ser Ala Gly Pro Glu Arg Asp Leu Ala
        35                  40                  45

Glu Leu Lys Ser Ser Leu Val Asn Glu Ser Glu Gly Ala Ala Gly Ser
    50                  55                  60

Ala Gly Ile Pro Gly Val Pro Gly Ala Gly Ala Gly Ala Arg Gly Glu
65                  70                  75                  80

Ala Glu Ala Leu Gly Arg Glu His Arg Ala Gln Arg Leu Phe Pro Asp
                85                  90                  95

Lys Leu Pro Glu Pro Leu Glu Asp Gly Leu Lys Ala Pro Glu Cys Thr
            100                 105                 110

Ser Gly Met Tyr Lys Glu Thr Val Tyr Ser Ala Phe Asn Leu Leu Met
        115                 120                 125

His Tyr Pro Pro Pro Ser Gly Ala Gly Gln His Pro Gln Pro Gln Pro
    130                 135                 140

Pro Leu His Lys Ala Asn Gln Pro Pro His Gly Val Pro Gln Leu Ser
145                 150                 155                 160

Leu Tyr Glu His Phe Asn Ser Pro His Pro Thr Pro Ala Pro Ala Asp
                165                 170                 175

Ile Ser Gln Lys Gln Val His Arg Pro Leu Gln Thr Pro Asp Leu Ser
            180                 185                 190

Gly Phe Tyr Ser Leu Thr Ser Gly Ser Met Gly Gln Leu Pro His Thr
        195                 200                 205

Val Ser Trp Pro Ser Pro Pro Leu Tyr Pro Leu Ser Pro Ser Cys Gly
    210                 215                 220

Tyr Arg Gln His Phe Pro Ala Pro Thr Ala Ala Pro Gly Ala Pro Tyr
225                 230                 235                 240

Pro Arg Phe Thr His Pro Ser Leu Met Leu Gly Ser Gly Val Pro Gly
                245                 250                 255

His Pro Ala Ala Ile Pro His Pro Ala Ile Val Pro Pro Ser Gly Lys
            260                 265                 270

Gln Glu Leu Gln Pro Phe Asp Arg Asn Leu Lys Thr Gln Ala Glu Ser
        275                 280                 285

Lys Ala Glu Lys Glu Ala Lys Lys Pro Thr Ile Lys Lys Pro Leu Asn
    290                 295                 300

Ala Phe Met Leu Tyr Met Lys Glu Met Arg Ala Lys Val Ile Ala Glu
305                 310                 315                 320

Cys Thr Leu Lys Glu Ser Ala Ala Ile Asn Gln Ile Leu Gly Arg Arg
                325                 330                 335

Trp His Ala Leu Ser Arg Glu Glu Gln Ala Lys Tyr Tyr Glu Leu Ala
            340                 345                 350

Arg Lys Glu Arg Gln Leu His Met Gln Leu Tyr Pro Gly Trp Ser Ala
        355                 360                 365

Arg Asp Asn Tyr Gly Lys Lys Lys Arg Arg Ser Arg Glu Lys His Gln
    370                 375                 380
```

-continued

Glu Ser Thr Thr Gly Gly Lys Arg Asn Ala Phe Gly Thr Tyr Pro Glu
385 390 395 400

Lys Ala Ala Ala Pro Ala Pro Phe Leu Pro Met Thr Val Leu Ala Ala
405 410 415

Pro Gly Pro Gln Leu Pro Arg Thr His Pro His Thr Ile Cys Cys Pro
420 425 430

Ala Ser Pro Gln Asn Cys Leu Leu Ala Leu Arg Ser Arg His Leu His
435 440 445

Pro Gln Val Ser Pro Leu Leu Ser Ala Ser Gln Pro Gln Gly Pro His
450 455 460

Arg Pro Pro Ala Ala Pro Cys Arg Ala His Arg Tyr Ser Asn Arg Asn
465 470 475 480

Leu Arg Asp Arg Trp Pro
485

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 511 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: None

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Met Pro Gln Leu Asp Ser Gly Gly Gly Gly Ala Gly Gly Gly Asp Asp
1 5 10 15

Leu Gly Ala Pro Asp Glu Leu Leu Ala Phe Gln Asp Glu Gly Glu Glu
20 25 30

Gln Asp Asp Lys Ser Arg Asp Ser Ala Gly Pro Glu Arg Asp Leu Ala
35 40 45

Glu Leu Lys Ser Ser Leu Val Asn Glu Ser Glu Gly Ala Ala Gly Ser
50 55 60

Ala Gly Ile Pro Gly Val Pro Gly Ala Gly Ala Gly Ala Arg Gly Glu
65 70 75 80

Ala Glu Ala Leu Gly Arg Glu His Arg Ala Gln Arg Leu Phe Pro Asp
85 90 95

Lys Leu Pro Glu Pro Leu Glu Asp Gly Leu Lys Ala Pro Glu Cys Thr
100 105 110

Ser Gly Met Tyr Lys Glu Thr Val Tyr Ser Ala Phe Asn Leu Leu Met
115 120 125

His Tyr Pro Pro Pro Ser Gly Ala Gly Gln His Pro Gln Pro Gln Pro
130 135 140

Pro Leu His Lys Ala Asn Gln Pro Pro His Gly Val Pro Gln Leu Ser
145 150 155 160

Leu Tyr Glu His Phe Asn Ser Pro His Pro Thr Pro Ala Pro Ala Asp
165 170 175

Ile Ser Gln Lys Gln Val His Arg Pro Leu Gln Thr Pro Asp Leu Ser
180 185 190

Gly Phe Tyr Ser Leu Thr Ser Gly Ser Met Gly Gln Leu Pro His Thr
195 200 205

Val Ser Trp Pro Ser Pro Pro Leu Tyr Pro Leu Ser Pro Ser Cys Gly
210 215 220

Tyr Arg Gln His Phe Pro Ala Pro Thr Ala Ala Pro Gly Ala Pro Tyr
225 230 235 240

-continued

Pro Arg Phe Thr His Pro Ser Leu Met Leu Gly Ser Gly Val Pro Gly
245 250 255

His Pro Ala Ala Ile Pro His Pro Ala Ile Val Pro Pro Ser Gly Lys
260 265 270

Gln Glu Leu Gln Pro Phe Asp Arg Asn Leu Lys Thr Gln Ala Glu Ser
275 280 285

Lys Ala Glu Lys Glu Ala Lys Lys Pro Thr Ile Lys Lys Pro Leu Asn
290 295 300

Ala Phe Met Leu Tyr Met Lys Glu Met Arg Ala Lys Val Ile Ala Glu
305 310 315 320

Cys Thr Leu Lys Glu Ser Ala Ala Ile Asn Gln Ile Leu Gly Arg Arg
325 330 335

Trp His Ala Leu Ser Arg Glu Glu Gln Ala Lys Tyr Tyr Glu Leu Ala
340 345 350

Arg Lys Glu Arg Gln Leu His Met Gln Leu Tyr Pro Gly Trp Ser Ala
355 360 365

Arg Asp Asn Tyr Gly Lys Lys Lys Arg Arg Ser Arg Glu Lys His Gln
370 375 380

Glu Ser Thr Thr Asp Pro Gly Ser Pro Lys Lys Cys Arg Ala Arg Phe
385 390 395 400

Gly Leu Asn Gln Gln Thr Asp Trp Cys Gly Pro Cys Arg Arg Lys Lys
405 410 415

Lys Cys Ile Arg Tyr Leu Pro Gly Glu Gly Arg Cys Pro Ser Pro Val
420 425 430

Pro Ser Asp Asp Ser Ala Leu Gly Cys Pro Gly Ser Pro Ala Pro Gln
435 440 445

Asp Ser Pro Ser Tyr His Leu Leu Pro Arg Phe Pro Thr Glu Leu Leu
450 455 460

Thr Ser Pro Ala Glu Pro Ala Pro Thr Ser Pro Gly Leu Ser Thr Ala
465 470 475 480

Leu Ser Leu Pro Thr Pro Gly Pro Pro Gln Ala Pro Arg Ser Thr Leu
485 490 495

Gln Ser Thr Gln Val Gln Gln Gln Glu Ser Gln Arg Gln Val Ala
500 505 510

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: None

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Ser Tyr Leu Asp Ser Gly Ile His Ser Gly Ala Thr Thr Thr Ala Pro
1 5 10 15

Ser Leu Ser Gly
20

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: None

-continued

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Ser Tyr Leu Gly Asp Ser Gly Ile His Ser Gly Ala Val Thr Gln Val
1 5 10 15

Pro Ser Leu Ser Gly
20

We claim:

1. A method of identifying candidate drugs for use in Familial Adenomatous Polyposis (FAP) patients, patients with Adenomatous Polyposis Coli (APC) or β-catenin mutations, or patients with increased risk of developing colorectal cancer, comprising the steps of:

contacting a cell having no wild-type APC or a mutant β-catenin with a test compound, wherein said cell comprises a TCF-responsive reporter gene;

measuring transcription of the Tcf-responsive reporter gene in said cell, wherein a test compound which inhibits the transcription of the reporter gene in said cell is a candidate drug for cancer therapy.

2. The method of claim 1 wherein the cell produces an APC protein defective in β-catenin binding or regulation.

3. The method of claim 1 wherein the cell produces a β-catenin protein which is super-active, or which is defective in APC binding or resistant to APC regulation.

4. The method of claim 1 wherein the cell produces no detectable APC protein.

5. A method of identifying candidate drugs for use in FAP patients, patients with APC or β-catenin mutations, or patients with increased risk of developing colorectal cancer, comprising the steps of:

contacting a Tcf-responsive reporter gene with a test compound under conditions in which the reporter gene is transcribed in the absence of the test compound; and

measuring transcription of the Tcf-responsive reporter gene; wherein a test compound which inhibits said transcription is a candidate drug for cancer therapy.

6. The method of claim 5 wherein the step of contacting is performed in the presence of a lysate of a cell which has no wild-type APC.

7. The method of claim 5 wherein the step of contacting is performed in the presence of a lysate of a cell which has a mutant β-catenin defective in APC binding or resistant to APC regulation or which is super-active.

8. The method of claim 6 wherein the cell produces an APC protein defective in β-catenin binding or regulation.

9. A method of identifying candidate drugs for use in PAP patients or patients with increased risk of developing colorectal cancer, comprising the steps of:

contacting a test compound with β-catenin and Tcf-4 under conditions in which β-catenin and Tcf-4 bind to each other; and

determining whether the test compound inhibits the binding of β-catenin and Tcf-4, a test compound which inhibits the binding being a candidate for cancer therapy or prophylaxis.

* * * * *

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 5,851,775

DATED : December 22, 1998

INVENTOR(S) : Barker et al.

It is certified that error appears in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Column 3, line 17, change "FIG. 1" to --FIGS. 1A, 1B and 1C--.

Column 3, line 29, change "FIG. 2" to --FIGS. 2A, 2B and 2C--.

Column 3, line 43, change "FIGS. 3A, 3B" to --FIGS. 3A, 3B and 3C--.

Column 4, line 44, change "FIGS. 7A, 7B and 7C" to --FIGS. 7A and 7B--.

Column 5, line 6, change "B" to --8B--.

Signed and Sealed this

Thirteenth Day of July, 1999

Attest:

Attesting Officer

Q. TODD DICKINSON

Acting Commissioner of Patents and Trademarks

FEBS Letters 356 (1994) 130–136

FEBS 14864

Tyrosine phosphorylation of the MUC1 breast cancer membrane proteins: Cytokine receptor-like molecules

Sheila Zrihan-Licht, Amos Baruch, Orna Elroy-Stein, Iafa Keydar, Daniel H. Wreschner*

Department of Cell Research and Immunology, The George S. Wise Faculty of Life Sciences, Tel Aviv University, Ramat Aviv, Israel

Received 12 September 1994; revised version received 27 October 1994

Abstract Phosphorylation on tyrosine residues is a key step in signal transduction pathways mediated by membrane proteins. Although it is known that human breast cancer tissue expresses at least 2 MUC1 type 1 membrane proteins (a polymorphic high molecular weight MUC1 glycoprotein that contains a variable number of tandem 20 amino acid repeat units, and the MUC1/Y protein that is not polymorphic and is lacking this repeat array) their function in the development of human breast cancer has remained elusive. Here it is shown that these MUC1 proteins are extensively phosphorylated, that phosphorylation occurs primarily on tyrosine residues and that following phosphorylation the MUC1 proteins may potentially interact with SH2 domain-containing proteins and thereby initiate a signal transduction cascade. As with cytokine receptors, the MUC1 proteins do not harbor intrinsic tyrosine kinase activity yet are tyrosine phosphorylated and the MUC1/Y protein participates in a cell surface heteromeric complex whose formation is mediated by two cytoplasmically located MUC1 cysteine residues. Furthermore, the MUC1/Y protein demonstrates sequence similarity with sequences present in cytokine receptors that are known to be involved in ligand binding. Our results demonstrate that the two MUC1 isoforms are both likely to function in signal transduction pathways and to be intimately linked to the oncogenetic process and suggest that the MUC1/Y protein may act in a similar fashion to cytokine receptors.

Key words: Breast cancer; Tyrosine phosphorylation; Receptor; MUC1

1. Introduction

Protein products of the MUC1 gene are expressed at high levels in adenocarcinomas and especially in human breast cancer tissue [1–7] and disease status in breast cancer patients is routinely assessed by monitoring the serum levels of circulating MUC1 proteins (variously referred to as episialin, H23Ag, ETA – epithelial tumor antigen, PEM – polymorphic epithelial mucin, EMA – epithelial membrane antigen, CA15–3, MCA – mammary carcinoma antigen, etc.). Molecular studies, including cDNA and gene cloning [8–14], have elucidated many properties of the MUC1 proteins. One of the MUC1 gene products is a polymorphic type 1 transmembrane molecule that consists of a large extracellular domain, a transmembrane domain and a 72 amino acid cytoplasmic tail (Fig. 1F, upper molecule). The genetic polymorphism derives from a tandem array of variable numbers of a highly conserved 20 amino acid repeat motif present within the extracellular domain. Soon after translation and prior to its translocation to the cell surface, this MUC1 protein (designated MUC1/REP) undergoes proteolytic cleavage in a region that is located 45 to 60 amino acids N-terminal to the transmembrane domain [15]. The two resulting protein molecules form a tight heterodimer complex that is composed of the large extracellular domain linked by non-covalent, SDS sensitive bonds to the much smaller (20–30 kDa) protein molecule containing the cytoplasmic and transmembrane domains [15]. Expression of the MUC1/REP protein in cell transfectants reduces cellular aggregation that is mediated by the highly glycosylated tandem repeat domain [16].

An additional novel MUC1 protein (designated MUC1/Y) has been recently characterized [8] that is devoid of the hallmark feature of MUC1, the tandem repeat array, yet retains the MUC1 N-terminal, transmembrane and cytoplasmic domains (Fig. 1F, lower molecule). The MUC1/Y protein is generated by a splicing mechanism that utilizes perfect alternative splice donor and splice acceptor sites located upstream and downstream to the tandem repeat array. Previous work demonstrated that the mature MUC1/Y protein has a molecular mass of between 42–45 kDa indicating that, in contrast to the cleaved MUC1/REP, it does not undergo proteolytic cleavage and is therefore continuous from its N-terminal extracellular domain through to its C-terminal cytoplasmic domain. In addition to the mature 42–45 kDa MUC1/Y protein, a precursor 33 kDa MUC1/Y protein that subsequently undergoes post-translational glycosylation modifications was also identified [8]. Significantly, both the MUC1/REP and MUC1/Y isoforms are highly expressed in human breast cancer tissue [1–8].

As the MUC1/REP and the novel MUC1/Y proteins are anchored at the cell surface and contain extracellular and common transmembrane and cytoplasmic domains they may both be involved in signal transduction processes.

Membrane proteins participating in signal transduction processes are in many cases modified by phosphorylation. It has not been known whether the MUC1 proteins are at all phosphorylated – indeed the MUC1 cytoplasmic domain does not contain any conserved sequence motifs known to exist in the catalytic domains of kinases in general or tyrosine kinases in particular, and is thus devoid of endogenous kinase activity and cannot undergo autophosphorylation. Similarly, cytokine receptors are also devoid of intrinsic kinase activity but are in many instances transphosphorylated on tyrosine residues by cytoplasmic tyrosine kinases [17–21]. It is shown here, for the first time, that the MUC1 proteins are phosphorylated on tyrosine residues and that following phosphorylation they have the potential to interact with SH2 domain containing proteins and thereby initiate a signal transduction cascade. Furthermore we show that, as with cytokine receptors, the MUC1/Y isoform participates in a cell-surface heteromeric complex. Interestingly, the MUC1/Y protein demonstrates sequence similarity with sequences present in cytokine receptors that are known to be involved in ligand binding. Our results demonstrate that the two MUC1 isoforms are both likely to function in signal

*Corresponding author. Fax: (972) (3) 642 2046.

SSDI 0014-5793(94)01251-2

EXHIBIT 4

transduction pathways and to be intimately linked to the oncogenetic process and suggest that the MUC1/Y protein may act in a similar fashion to cytokine receptors.

2. Materials and methods

2.1. Cells

Stable transfectants were generated by co-transfecting an expression plasmid harboring either MUC1/REP or MUC1/Y cDNA with the neomycin plasmid (pSV2 neo) selection marker into HBL100 human mammary epithelial cells (MUC1/REP transfectants) or into 3T3 *ras* transformed fibroblasts (MUC1/Y transfectants). These cells were chosen as they showed the highest expression levels of the respective MUC1 isoforms. Transient transfectants were prepared using the T7/ encephalomyocarditis virus/vaccinia sysytem (described below) and the monkey epithelial cell line, BSC-1.

2.2. cDNA constructs

The generation of expression vectors harboring either the full-length transmembrane MUC1/REP or the novel MUC1/Y cDNA and driven by the HMG coenzyme A reductase promoter (expression vector pCL642) has been previously described [8].

2.3. Transient expression of the novel MUC1 protein in the T7/EMC/ vaccinia system

Transient expression of the novel MUC1/Y protein synthesized in a state as close as possible to the naturally occurring MUC1/Y protein, was accomplished using the T7/EMC/vaccinia hybrid expression system [8]. This system utilizes the bacteriophage T7 RNA polymerase which is encoded by a recombinant vaccinia virus to transcribe genes that are regulated by the T7 promoter in the cytoplasm of infected mammalian cells. The MUC1/Y cDNA was inserted into the pTM1 vector under the control of the T7 promoter and EMCV leader, and introduced into tissue cultured BSC-1 epithelial cells together with recombinant vaccinia virus which expresses T7 RNA polymerase [8].

2.4. Western blot analyses

Cell lysates were prepared by adding lysis buffer (50 mM NaCl, 20 mM Tris-HCl, pH 7.4, 100 μg/ml leupeptin and 0.5% Nonidet P-40) to cell pellets, followed by vortex mixing and sonication (3 times 10 second bursts using a Branson sonicator). Cell debris was removed by centrifugation at 10,000 rpm for 10 min. All procedures were performed at 4°C or on ice. Protein samples were denatured by boiling in SDS buffer containing mercaptoethanol and analyzed on SDS/acrylamide gels. The gel was electrotransferred for 3 h at 1 A to nitrocellulose filters that were then blocked in PBS containing 5% skimmed milk followed by incubation with the primary antibody. The filters were washed in PBS and then incubated with a secondary anti-rabbit (or anti-mouse antibody) conjugated to horseradish peroxidase followed by ECL (Amersham) detection.

2.5. Antibodies

Western blot analyses were performed with a polyclonal antibody (a kind gift from Dr. Sandra Gendler) directed against the oligopeptide SSLSYTNPAVAATSANL (amino acids 499 to 515, see [8] for amino acid numbering) which represents the C-terminal region of the MUC1 cytoplasmic domain. The polyclonal antibody was precleared by adsorption against glutaraldehyde insolubilized human serum.

2.6. Metabolic labelling of cells with phosphate

Cells were incubated overnight in low phosphate medium supplemented with 10% fetal calf serum that had been dialyzed against saline. The following day radioactive carrier-free inorganic phosphate was added to the cells and incubation continued for another 8 h. Thirty minutes prior to harvesting, the cells were treated with the tyrosine phosphatase inhibitors, sodium vanadate (200 μM) and hydrogen peroxide (200 μM).

2.7. Immunoprecipitations

Cell lysates prepared as described above were added to protein-A-agarose-antibody complexes and incubated for 2 h at 4°C. The immunocomplex was washed 3 times with cell lysis buffer and 2× SDS sample buffer was added.

3. Results and discussion

To investigate whether MUC1 is transphosphorylated, stable transfectants expressing either the MUC1/REP protein or the novel MUC1/Y protein were generated. Immunoblotting experiments with antibodies directed against the MUC1 cytoplasmic domain confirmed MUC1/REP (20–30 kDa immunoreactive proteins) and MUC1/Y (42–45 kDa immunoreactive proteins) expression in the respective transfectants (Fig. 1A). MUC1 expressing transfectants were incubated with radioactively labelled inorganic phosphate in the presence of the tyrosine phosphatase inhibitors, hydrogen peroxide and sodium vanadate [22,23], cell lysates were then prepared and subjected to immunoprecipitation with anticytoplasmic domain antibodies. The specifically immunoprecipitated proteins migrating with molecular masses of 20–30 kDa for the MUC1/REP protein (Fig. 1B, lane 2) and 42–45 kDa for the MUC1/Y protein (Fig. 1B, lane 6) were highly labelled, indicating that the MUC1 proteins had undergone extensive phosphorylation. Similarly the MUC1 proteins were also found to be phosphorylated in non-transfected human T47D breast cancer cells (data not shown). Due however to the considerably lower level of expression as compared to that in the MUC1 transfectants, the signal of the phosphorylated MUC1 proteins in the T47D cells was correspondingly lower and further work was thus conducted with the MUC1 transfectants.

The effect of the tyrosine phosphatase inhibitors on the levels of MUC1 phosphorylation was next investigated. In the absence of these inhibitors, MUC1 phosphorylation demonstrated a low yet signficant level of phosphorylation that in their presence was markedly enhanced (Fig. 1C,D) suggesting that phosphorylation of the MUC1 proteins occurs predominantly on tyrosine residues. Consistent with this, nonspecifically precipitated labelled proteins (Fig. 1C, open arrow at left of figure) showed no differential enhancement of phosphorylation following treatment of cells with the tyrosine phosphatase inhibitors.

A phosphoamino acid analysis performed on the labelled phosphorylated MUC1 proteins showed that phosphorylation had indeed primarily occurred on tyrosine residues (70–90% in different experiments), with much reduced levels of phosphoserine and undetectable levels of threonine phosphorylation (Fig. 1E (a)). This pattern of tyrosine phosphorylation was observed both for the MUC1/REP and MUC1/Y proteins. Further confirmation for tyrosine phosphorylation of the MUC1 proteins was obtained by probing immunoblots of immunoprecipitated MUC1 proteins with antiphosphotyrosine antibodies. This analysis (Fig. 1E (b)) clearly showed that the MUC1/Y protein is readily detected by antiphosphotyrosine antibodies following treatment of cells with tyrosine phosphatase inhibitors.

Three independent lines of evidence thus support the finding that the MUC1 proteins are phosphorylated on tyrosine residues: (i) increased levels of MUC1 phosphorylation following treatment of cells with tyrosine phosphatase inhibitors, (ii) a phosphoaminoacid analysis of the MUC1 proteins, and (iii) reactivity of phosphorylated MUC1 proteins with antiphosphotyrosine antibodies.

Interestingly, tyrosine residues are distributed in a markedly biased fashion within the MUC1 proteins – 7 out of 72 of the amino acids comprising the MUC1 cytoplasmic domain are

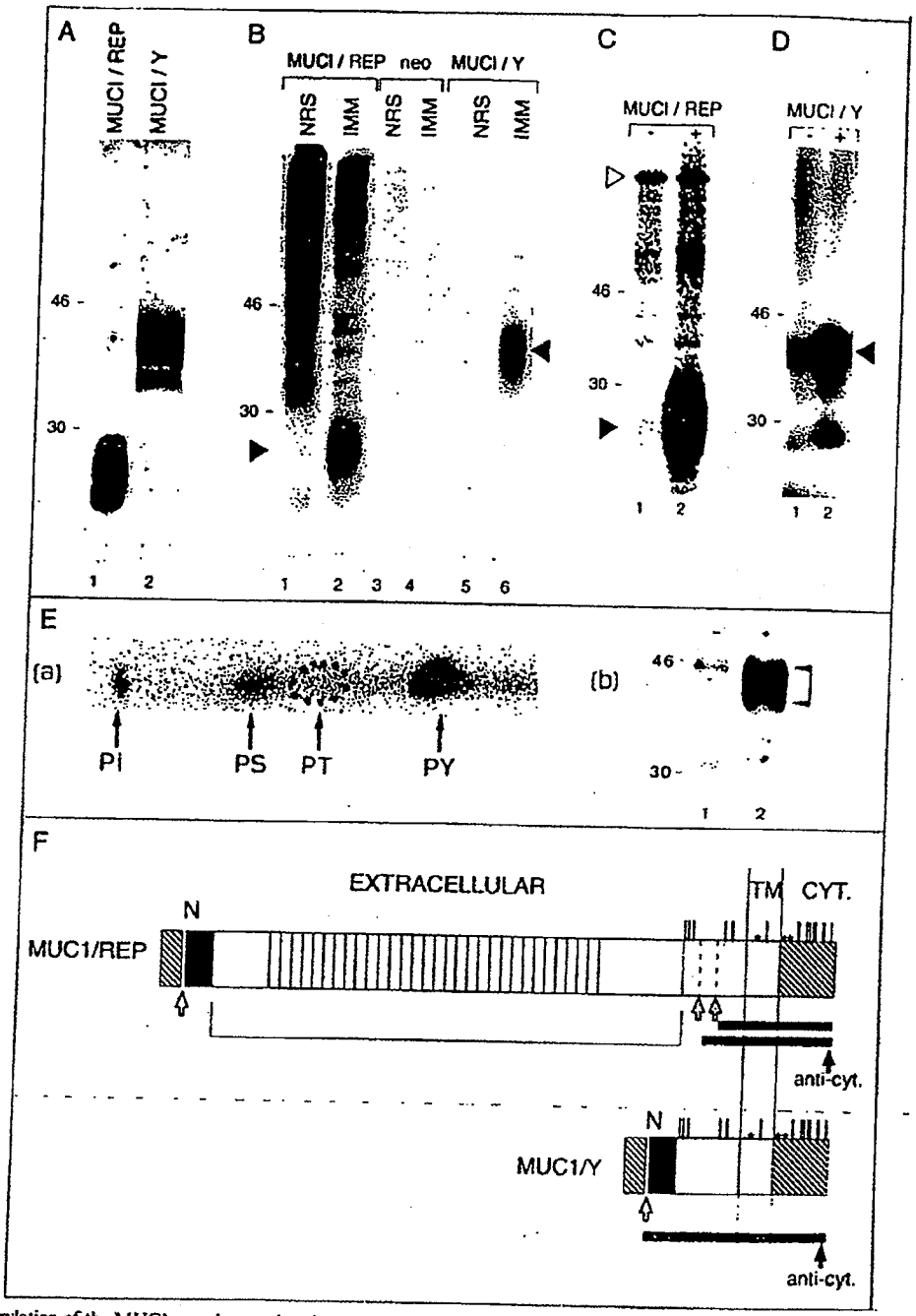


Fig. 1. Phosphorylation of the MUC1 proteins on phosphotyrosine residues. (A) MUC1/REP and MUC1/Y expression in stable transfectants. lysates were prepared from the MUC1/REP (lane 1) and MUC1/Y transfectants (lane 2) and the proteins resolved on SDS-polyacrylamide (gels, transferred to nitrocellulose and probed with polyclonal antisera directed against the MUC1 cytoplasmic domain, as described in secti Molecular size standards (in kilodaltons) are shown at left. (B) Phosphorylation of the MUC1 proteins. The MUC1/REP transfectants (lanes 1 2, MUC1/REP), control neomycin transfected ras 3T3 fibroblasts (lanes 3 and 4, neo) and MUC1/Y ras 3T3 fibroblast transfectants (lanes 5 6, MUC1/Y) were labelled with radioactive carrier-free inorganic phosphate (section 2). The cells were then harvested, briefly washed phosphate-buffered saline and cell lysates prepared that were then subjected to immunoprecipitation with either preimmune rabbit serum (NRS, l 1, 3 and 5) or anticytoplasmic domain antisera (IMM, lanes 2, 4 and 6). Precipitated proteins were resolved on SDS-polyacrylamide (10%) gels visualized by autoradiography. (C) and (D) Effect of tyrosine phosphatase inhibitors on MUC1/REP (C) and MUC1/Y (D) phosphorylat

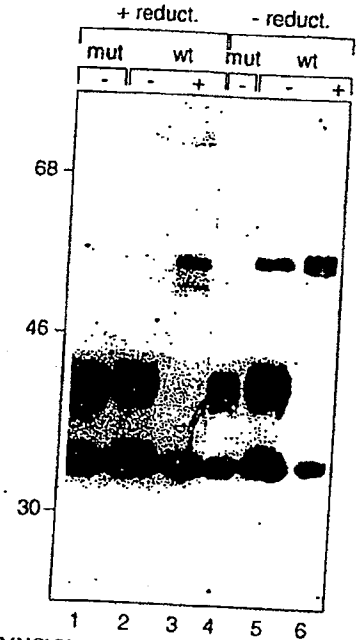


Fig. 2. The MUC1/Y protein appears in a cell surface complex, the formation of which is mediated via MUC1 cytoplasmic domain cysteine residues. Cell lysates were prepared from monkey BSC-1cells infected with recombinant vaccinia virus coding for T7 RNA polymerase and transfected with the pTM1 vaccinia expression vector harboring cDNA coding for either wild type MUC1/Y protein (wt, lanes 2, 3, 5 and 6) or mutant MUC1/Y protein (mut, lanes 1 and 4) in which the Cys-Gln-Cys (CQC) sequence had been mutated to Gly-Gln-Gly (see Fig. 1F for the location of the CQC sequence). The cell lysate proteins were resolved on SDS polyacrylamide (10%) gels under non-reducing (− reduct., lanes 4–6) or reducing (+ reduct., lanes 1–3) conditions, transferred to nitrocellulose and probed with polyclonal antisera directed against the MUC1 cytoplasmic domain. Proteins resolved in lanes 3 and 6 were derived from cells that prior to harvesting had been treated for 15 min with 15 mM EDAC crosslinking agent (obtained from Sigma). Detection of bound antibodies was performed as described in Fig. 1A.

tyrosine residues (Fig. 1F). The transmembrane domain contains one tyrosine residue and a further 5 tyrosine residues appear within the 92 amino acids N-terminally adjacent to the transmembrane domain – no additional tyrosine residues appear within the MUC1 proteins. The MUC1 amino acid sequence also reveals a marked similarity between tyrosine containing sequences located within the MUC1 cytoplasmic domain and phosphotyrosine containing peptide sequences that are *postulated* to specifically interact with SH2 domain containing proteins [26]. It should be emphasized that these sites represent only *presumptive* docking sites for SH2 domain containing proteins; it is nonetheless striking that the 72 amino acid MUC1 cytoplasmic domain contains no less than 3 such possible sites. For example, the most preferred sequence for interaction with phospholipase C γ 1 is pTyr-Val-Iso-Pro (pYVIP) and a very similar sequence [Tyr-Val-Pro-Pro (YVPP)] appears in the cytoplasmic domain of the MUC1 protein. Additionally the sequence pTyr-Glu-Glu-Val (pYEEV) which is identical to a sequence that appears within the mouse MUC1 cytoplasmic domain, has been shown to be one of the most preferred sequences for interaction with a number of SH2 domain containing cytoplasmic tyrosine kinases [26] and a potential GRB-2 binding site (pYXNX) also appears in the MUC1 cytoplasmic domain. That the MUC1 cytoplasmic domain *has the potential* to interact with SH2 domain containing proteins has been experimentally demonstrated by the binding of in-vitro tyrosine phosphorylated MUC1 cytoplasmic domain to the src SH2 domain, the SH2 domain derived from the N-terminal part of phospholipase C and to the GRB-2 protein (data not shown); no binding was observed to the SH2 domain derived from the C-terminal portion of p85 phosphatidyl inositol (PI) 3′ kinase. One should bear in mind that in-vitro tyrosine phosphorylated MUC1 cytoplasmic domain may not faithfully reflect the tyrosine phosphorylation state of this protein within the cell; experiments investigating the actual association of the MUC1 protein with SH2 domain second messenger proteins in vivo are presently being conducted. Nonetheless, the analyses described above do indicate that the tyrosine phosphorylated MUC1 protein certainly has the potential to participate in such interactions.

Fig. 1 (*continued*).

MUC1/REP or MUC1/Y transfectants were labelled with radioactive carrier-free inorganic phosphate as described above and either treated with sodium vanadate and hydrogen peroxide (lane 2, +) or not treated (lane1, −). Proteins immunoprecipitated with anticytoplasmic domain antibodies were visualized as above. Non-specifically immunoprecipitated proteins are indicated in (C) by the open arrow; the closed arrows indicate the specifically immunoprecipitated MUC1 proteins. (E) (a) Phosphoamino acid analysis. The immunoprecipitated phosphorylated MUC1 proteins (from lane 6 in (B)) were isolated from SDS-acrylamide (10%) gel and hydrolyzed in 6M HCl at 110°C for 1 h. Labelled phosphoaminoacids (with added unlabelled internal phosphoamino acid markers) were analyzed by thin-layer high voltage electrophoresis followed by Phosphorimager analysis. The position of migration of phosphoserine, phosphothreonine and phosphotyrosine are indicated by PS, PT and PY respectively and inorganic phosphate is shown by P_i. Phosphoamino acid analyses performed on the phosphorylated MUC1 cleavage products (lanes 1 and 2 in (C)) and the phosphorylated MUC1/Y proteins [lanes 1 and 2 in (D)] gave similar results (data not shown). (b) Reactivity of MUC1/Y protein with anti-phosphotyrosine antibodies. MUC1/Y ras 3T3 fibroblast transfectants were either treated with sodium vanadate and hydrogen peroxide (lane 2) or not treated (lane 1). Cell lysates were prepared and proteins immunoprecipitated with anticytoplasmic domain antibodies were resolved on SDS-polyacrylamide (10%) gels followed by immunoblotting and probing with antiphosphotyrosine antibodies (PY20 and PY69). The arrowed region indicates the tyrosine phosphorylated MUC1/Y protein. (F) Scheme depicting the repeat array containing MUC1 protein (upper molecule) and the novel MUC1/Y protein (lower molecule) the novel MUC1/Y form is generated by an alternative splicing event that deletes the central tandem repeat array (compare upper and lower molecules). The location of tyrosine and cysteine residues are indicated above the rectangles by vertical lines and asterisks, respectively. Both MUC1 forms contain a hydrophobic N-terminal signal sequence (slashed box at left of figure) that is co-translationally cleaved (arrow at left of figure). This is followed by the tandem repeat array (upper molecule) that is illustrated by the block of closely spaced vertical lines. The highly hydrophobic 28 amino acid stretch constituting the transmembrane domain (TM) is shown at the C-terminal end of both MUC1 proteins, followed by the cytoplasmic domain (CYT). The region comprising the proteolytic cleavage site [15] of the repeat array containing MUC1 protein (upper molecule) is indicated by the 2 vertical dotted arrows just N-terminal to the transmembrane domain. The regions recognized by the anti-cytoplasmic domain (anti-...) antibodies are indicated.

Considering that within the cell less than 1% of total protein phosphorylation occurs on tyrosine residues, and that when phosphorylated, phosphotyrosine residues play a pivotal role in signal transduction processes, both the extensive tyrosine phosphorylation of the MUC1 proteins (Fig. 1) and the biased distribution of MUC1 tyrosine containing sequences (Fig. 1) that can potentially interact with SH2 domain containing proteins, are especially significant findings.

In an analogous fashion to a number of cytokine receptors [17–21,30], the MUC1 proteins, as shown above, do not have intrinsic tyrosine kinase activity yet are phosphorylated on tyrosine residues. As cytokine receptors are displayed on the cell surface as heteromeric protein complexes [17–21], we wished to investigate whether the MUC1 proteins also form cell surface heteromeric complexes. Integrity of the MUC1 extracellular domain is likely to be essential for binding of a putative ligand, and these studies therefore concentrated on the MUC1/Y isoform. Addition to MUC1/Y expressing transfectants of a crosslinking agent that does not penetrate the cell membrane lead to a substantial decline in the level of the mature cell-surface located 42–45 kDa MUC1/Y protein and concomitantly to the appearance of a new 60 kDa band (Fig. 2, lanes 3 and 6), thereby demonstrating that the MUC1/Y protein is complexed with other cell-surface located molecules. Notably, the cytoplasmically located precursor 33 kDa MUC1/Y protein was not affected by treatment with the crosslinking agent (Fig. 2, lanes 3 and 6), indicating that complex formation involved only the mature cell surface located MUC1/Y protein. It should be noted that in subsequent cross-linking experiments the 60 kDa MUC1/Y complex was consistently observed although the levels of the cell-surface located 42–45 kDa MUC1/Y protein were somewhat higher than those observed in the experiment described above – the reason for this variability is unknown. Furthermore, gel analysis of the MUC1/Y protein from cells not treated with the crosslinking agent showed, under non-reducing conditions and in the presence of iodoacetamide, a prominent 60 kDa band (Fig. 2, lane 5) that migrated to an identical position as the cross-linked 60 kDa MUC1/Y protein. This band was not observed under reducing conditions (compare Fig. 2, lanes 2 and 5), indicating that MUC1/Y complex formation is likely to be mediated, at least in part, by cysteine residues that form reducible disulfide bridges. The recently described interferon α/β cytokine receptor [31] has also been shown to form disulfide-linked dimer complexes – the cysteine residues within this cytokine receptor that are responsible for dimerization have, however, not been identified. The MUC1 protein contains only 3 cysteines – one cysteine residue appears within the transmembrane domain and the remaining two are located in a Cys-Gln-Cys tripeptide just C-terminal to the transmembrane domain (see Fig. 1F for location of cysteine residues within the complete MUC1 protein). An identical cytoplasmically located Cys-Gln-Cys tripeptide sequence has been previously shown to mediate complex formation of the cell surface CD4 molecule [32], and we thus investigated whether the MUC1 Cys-Gln-Cys sequence may play a similar role in MUC1/Y complex formation. Transfectants were generated that expressed a mutant form of MUC1/Y in which the Cys-Gln-Cys sequence had been mutated to Gly-Gln-Gly. These mutants expressed the 42–45 kDa MUC1/Y protein that migrated to an identical position as the wild type MUC1/Y protein (Fig. 2, lane 1). However, in marked contrast to the wild type MUC1/Y protein, the mutant MUC1/Y protein did not form, under non-reducing conditions, the 60 kDa complex (Fig. 2, lane 4). Additional experiments involving treatment of the mutant MUC1/Y transfectant with the crosslinking agent demonstrated only very low levels of the MUC1/Y 60 kDa complex (data not shown). This indicates that formation of the MUC1/Y 60 kDa complex is primarily mediated by the two cysteine residues present in the Cys-Gln-Cys tripeptide which form disulfide bridges, and that non-covalent protein–protein interactions may only play a minor role in complex formation. Although the molecular mass of the 60 kDa complex suggests that it is a disulfide linked heterodimer, we cannot rule out at this stage the possibility of an anomalously migrating homodimer or even a complex of more than 2 protein molecules.

It is thereby demonstrated that, as in the case of cytokine receptors, the MUC1/Y protein is also presented at the cell surface complexed to other membrane proteins. Furthermore, the MUC1 amino acid sequence reveals striking similarities to cytokine receptor sequences that are known to participate in ligand binding [33–36] (Fig. 3). For example, when considering only identical amino acid residues and Ser-Thr substitutions the

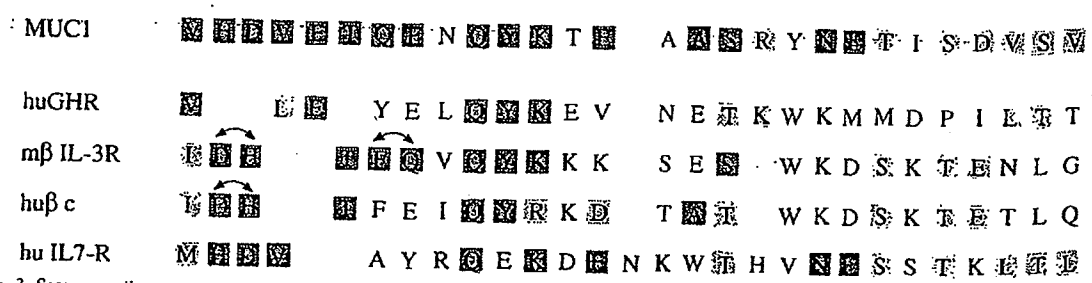


Fig. 3. Sequence alignment of residues in the MUC1 extracellular domain with the predicted ligand binding domains of cytokine receptors. The MUC1/Y extracellular domain amino acid sequence starting at amino acid number 114 [8] is compared to the predicted loop between the C-terminal B and C strands in the proposed double β-barrel structural models of the human interleukin 7 receptor (huIL-7R), β chain mouse IL-3 receptor (mβIL-3R), common β subunit of human GM-CSF (granulocyte-macrophage colony stimulating factor), IL-3 and IL-5 receptors (huβcR) and the human growth hormone receptor (huGHR). Identical amino acids are heavily shaded and conservatively substituted amino acids are lightly shaded. Note that the His-Asp (HD) dipeptide sequence that is proximal to the N-terminal of both the IL-7 receptor and MUC1 sequences appears as Asp-His (DH) both in the β chain mouse IL-3 receptor (mβIL-3R) and in the common β subunit of human GM-CSF (granulocyte-macrophage colony stimulating factor), IL-3 and IL-5 receptors (huβcR), and the Gln-Phe (QF) dipeptide that is present in the MUC1 sequence appears in the β chain mouse IL-3 receptor (m IL-3R) as Phe-Gln (FQ).

human IL-7 receptor shows 44% homology with MUC1 extracellular domain sequences over a stretch of 27 amino acids that span the ligand binding site- if one also includes in this homology comparison valine to leucine and valine to methionine substitutions, the extent of homology increases to 55%. Significantly this homology maps in close proximity to the region where proteolytic cleavage occurs in the MUC1/REP protein, suggesting that integrity of this site in the MUC1/Y protein may be of prime importance for both ligand binding and signal transmission and that different mechanisms may be responsible for activation of the two MUC1 isoforms. The MUC1/Y protein, however, contains neither the conserved extracellular domain cysteine residues nor the Trp-Ser-Xxx-Trp-Ser, motif that are characteristic of many cytokine receptors [20]- it is notable that this latter motif appears in the human growth hormone receptor as Tyr-Gly-Glu-Phe-Ser and not as the canonical WSXWS. It is therefore interesting that a Phe-Ser-Xxx-Xxx-Ser motif (Phe-Ser-Ala-Gln-Ser) does appear in the MUC1 sequence just N-terminal to the transmembrane domain, at an identical location to the WSXWS motif, seen in cytokine receptors.

Taken together, these data demonstrate that the MUC1 proteins participate in signal transduction and that the MUC1/Y protein may act as a cytokine receptor-like molecule.

It has been shown here for the first time that the MUC1 proteins are extensively phosphorylated, that phosphorylation occurs on tyrosine residues and that following phosphorylation the MUC1 proteins may potentially interact with SH2 domain containing proteins and thereby initiate a signal transduction cascade. As with cytokine receptors [17], the MUC1/Y protein does not harbor intrinsic tyrosine kinase activity yet is tyrosine phosphorylated, and participates in a cell-surface heteromeric complex – furthermore, the MUC1/Y protein demonstrates sequence similarity with sequences present in cytokine receptors that are known to be involved in ligand binding. Indeed, the striking enhancement of MUC1 phosphorylation (Fig. 1C,D) elicited by the tyrosine phosphatase inhibitors suggests the existence of regulatory mechanisms, such as ligand binding to the MUC1 extracellular domains or/and cellular redox potential changes [37] that may control MUC1 tyrosine phosphorylation levels by activating cytoplasmic kinases which subsequently transphosphorylate the MUC1 proteins. These features suggest that the MUC1/Y protein may act in a similar fashion to cytokine receptors and that following binding of an as yet unidentified ligand, undergo transphosphorylation mediated by cytoplasmic tyrosine kinases such as the Janus kinases [17,30].

Notwithstanding the fact that the two MUC1 isoforms have identical cytoplasmic domains, are phosphorylated on tyrosine residues and are both likely to participate in signal transduction processes, the difference in their extracellular domain structure and the cleavage of the MUC1/REP form as opposed to the integrity of MUC1/Y all argue against identical functions as well as activating mechanisms for the two isoforms.

We have previously shown that both the MUC1/REP and MUC1/Y proteins are highly expressed in human breast cancer tissue [8]. The elucidation of mechanisms that activate the cell-surface located MUC1 proteins, shown here to be intimately linked to signal transduction and oncogenetic processes, may lead to new modalities for the treatment of human breast cancer.

Acknowledgements: This work was supported in part by the Israel Cancer Association and the Naftali Foundation (D.H.W. and O.E.S.), Simko Chair for Breast Cancer Research, Federico Fund for Tel Aviv University, Barbara Friedman Fund and The Washington Friends of T.A.U. (I.K.).

References

[1] Ceriani, R.L., Peterson, J.A., Lee, J.Y., Moncada, R. and Blank, E.W. (1983) Som. Cell Genet. 9, 415–427.
[2] Hilkens, J., Buijs, F., Hilgers, J., Hageman, Ph., Calafat, J., Sonnenberg, A. and Van der Valk, M. (1984) Int. J. Cancer 34, 197–206.
[3] Sekine, H., Ohno, T. and Kufe, D.W. (1985) J. Immunol. 135, 3610–3615.
[4] Linsley, P.S., Kallestad, J.C. and Horn, D. J. Biol. Chem. 263, 8390–8397.
[5] Gendler, S.J., Taylor-Papadimitriou, J., Duhig, T., Rothbard, J. and Burchell, J.M. (1988) J. Biol. Chem. 263, 12820–12823.
[6] Keydar, I., Chou, C.S., Hareuveni, M., Tsarfaty, I., Sahar, E., Seltzer, G., Chaitchik, S. and Hizi, A. (1989) Proc. Natl. Acad. Sci. 86, 1362–1366.
[7] Devine, P.L. and McKenzie, I.F.C. (1992) BioEssays 14, 619–625.
[8] Zrihan-Licht, S., Vos, H.L., Baruch, A., Elroy-Stein, O., Sagiv, D., Keydar, I., Hilkens, J. and Wreschner, D.H. (1994) Eur. J. Biochem. 224, 787–795.
[9] Wreschner, D.H., Tsarfaty, I., Hareuveni, M., Zaretsky, J., Smorodinsky, N., Weiss, M., Horev, J., Kotkes, P., Zrihan, S., Jeltsch, J.M., Green, S., Lathe, R. and Keydar, I. In: Breast Cancer: Progress in Biology, Clinical Management and Prevention, M.A. Rich, J.C. Hager and I. Keydar, Eds. (Kluwer Academic Publishers, Boston) pp. 41–59. 1989.
[10] Wreschner, D.H., Hareuveni, M., Tsarfaty, I., Smorodinsky, N., Horev, J., Zaretsky, J., Kotkes, P., Weiss, M., Lathe, R., Dion, A.S. and Keydar, I. (1990) Eur. J. Biochem. 189, 463–473.
[11] Ligtenberg, M.J.L., Vos, H.L., Genissen, A.M.C. and Hilkens, J. (1990) J. Biol. Chem. 265, 5573–5578.
[12] Gendler, S.J., Lancaster, S.J., Taylor-Papadimitriou, J., Duhig, T., Peat, N., Burchell, J., Pemberton, L., Lalani, E.-N. and Wilson, D.J. (1990) Biol Chem. 265, 15286–15293.
[13] Lan, M.S., Batra, S.K., Qi, W.-N., Metzgar, R.S. and Hollingsworth, M.A. (1990) J. Biol. Chem. 265, 15294–15299.
[14] Abe, M., Kufe, D. Abe, M. and Kufe, D. (1993) Proc. Natl. Acad. Sci. USA 90, 282–286.
[15] Ligtenberg, M.J.L., Kruijshaar, L., Buijs, F., van Meijer, M., Litvinov, S.V. and Hilkens, J. (1992) J. Biol. Chem. 267, 6171–6177.
[16] Ligtenberg, M.J.L., Buijs, F., Vos, H.L. and Hilkens, J. (1992) Cancer Res. 52, 2318–2324.
[17] Stahl, N. and Yancopoulos, G.D. Cell 74, 587–590.
[18] Hatakeyama, M., Kono, T., Kobayashi, N., Levin, S.D., Perlmutter, R.M. and Taniguchi, T. (1991) Science 252, 1523–1528.
[19] Wittuhn, B.A., Silvennoinen, O., Miura, O., Siew Lai, K., Cwik, C., Liu, E.T. and Ihle, J.N. (1994) Nature 370, 153–157.
[20] Miyajima, A., Kitamura, T., Harada, N., Yokota, T., Arai, K. (1992) Annu. Rev. Immunol. 10, 295–331
[21] Ziemiecki, A., Harpur, A.G., Wilks, A.F. (1994) Trends Cell Biol. 4, 207–212.
[22] Heffetz, D., Bushkin, I., Dror, R., Zick, Y. (1990) J. Biol. Chem. 265, 2896–2902
[23] Hecht, D. and Zick, Y. (1992) Biochem. Biophys. Res. Commun. 188, 773–778.
[24] Koch, A.C., Anderson, D., Moran, M.F., Ellis, C., Pawson, T. (1991) Science 252, 668–674.
[25] Pawson, T. and Schlessinger, J. (1993) J. Curr. Biol. 3, 434–442.
[26] Songyang, Z., Shoelson, S.E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W.G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R.J., Neel, B.G., Birge, R.B., Fajardo, J.E., Chou, M.M., Hanafusa, H., Schaffhausen, B. and Cantley, L.C. (1993) Cell 72, 767–778.

[27] Buday, L. and Downward, J. (1993) Cell 73, 611–620.
[28] Skolnik, E.Y., Lee, C.H., Batzer, A.M., Vicentini, L.M., Zhou, M., Daly, R., Myers, M.J., Backer, J.M., Ullrich, A., White, M.F. and Schlessinger, J. (1993) EMBO J. 12, 1929–1936.
[29] Puil, L., Liu, J., Gish, G., Mbamalu, G., Bowtell, D., Pelicci, P.G., Arlinghaus, R. and Pawson, T. (1994) EMBO J. 13, 764–773.
[30] Darnell, J.E., Kerr, I.M. and Stark, G.R. (1994) Science 264, 1415–1421.
[31] Novick, D., Cohen, B. and Rubinstein, M. (1994) Cell 77, 391–400.
[32] Turner, J.M., Brodsky, M.H., Irving, B.A., Levin, S.D., Perlmutter, R.M. and Littman, D.R. (1990) Cell 60, 755–765.
[33] deVos, A.M., Ultsch, M. and Kossiakoff, A.A. (1992) Science 255, 306–312.
[34] Imler, J.-L., Miyajima, A. and Zurawski, G. (1992) EMBO J. 11, 2047–2053
[35] Wang, H.-M., Ogorochi, T., Arai, K.-i. and Miyajima, A.J. (1992) Biol. Chem. 267, 979–983.
[36] Lock, P., Metcalf, D. and Nicola, N. (1994) Proc. Natl. Acad. Sci. USA 91, 252–256.
[37] Bauskin, A.R., Alkalay, I. and Ben-Neriah, Y. (1991) Cell 66, 685–696.